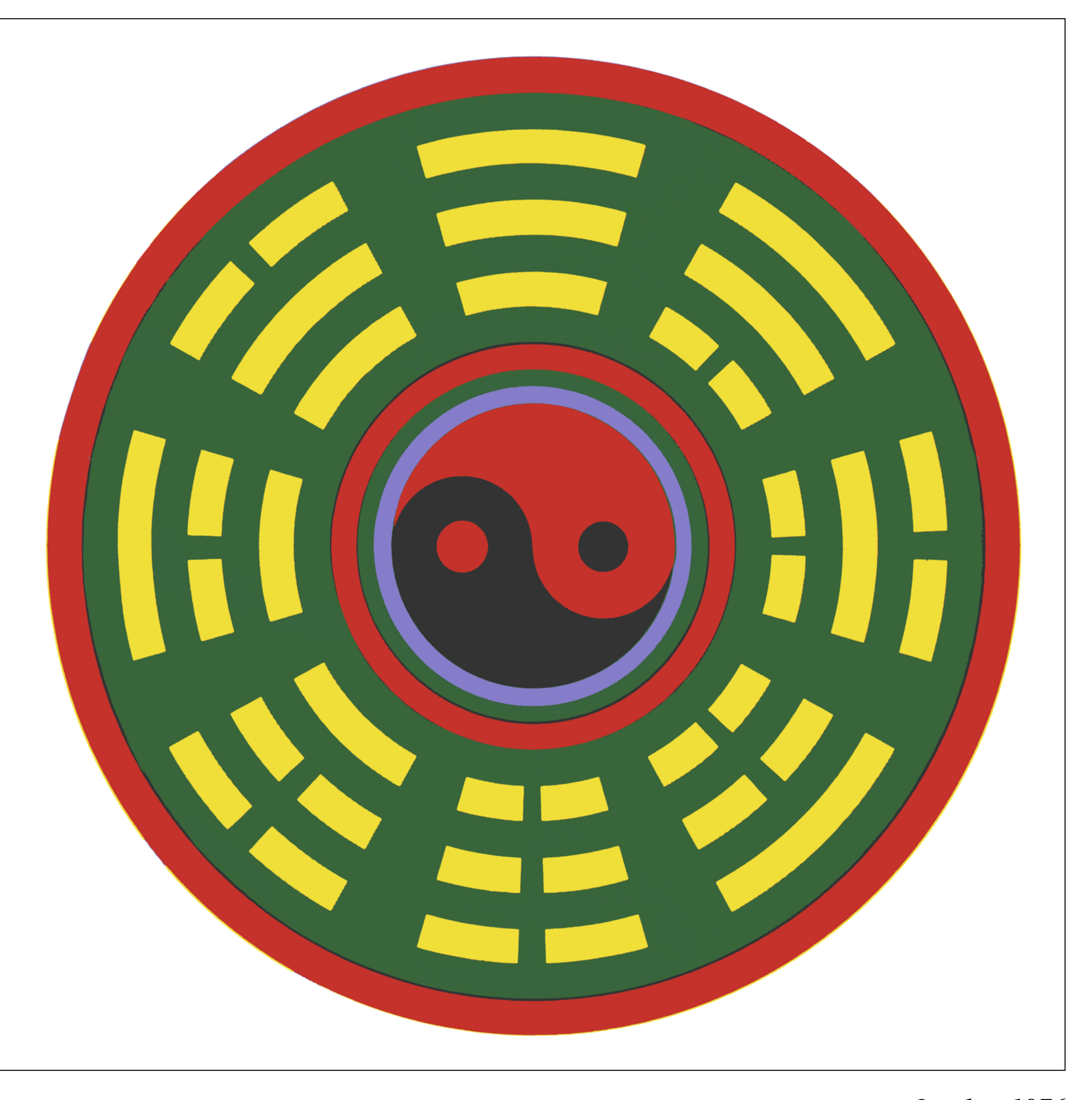
INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



LAS MATEMATICAS DEL "I CHING"

Octubre 1976

140 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 1
Octubre 1976

INVESTIGACION Y
CIENCIA
Edición en español de SCIENTIFIC AMERICAN

Colaboradores de este número:

Traducción:

Luis Mas: *¿Expansión indefinida del universo?;* Luis Alonso: *La búsqueda de nuevas familias de partículas elementales;* Juan Antonio Vela: *Las proteínas cromosómicas y la regulación de los genes;* Antonio Subías: *El problema del cáncer;* María José Riu: *El surgimiento de una clase mercantil maya;* Francisco Serra: *Sistemas robot;* Xavier Llimona: *Vinos, vides y clima;* Miguel Llobera: *Taller y laboratorio;* Teresa Riera: *Juegos matemáticos.*

Asunción López Collados: datos complementarios para el artículo *Vinos, vides y clima.*

Ciencia y sociedad:
Antonio Romañá.

Libros:
Luis Alonso y Philip Morrison



LA PORTADA

La ilustración de la cubierta se basa en un amuleto chino en el que se disponen los ocho trigramas del *I Ching* en torno al *t'ai chi t'u,* viejo símbolo chino del yin y del yang (véase "Juegos matemáticos", página 111). Esta distribución tradicional, atribuida a Fu Hsi, mítico fundador de la primera dinastía china, coloca los trigramas de forma tal que los pares opuestos sean complementarios en su sentido matemático y en su significación simbólica. El trigrama Ch'ien, en el centro de la parte superior, representa al padre y al cielo; su antagónico inferior es K'un, que simboliza a la madre y a la tierra. De igual forma, Tui (el lago, la menor de las hijas) se opone a Ken (la montaña, el menor de los hijos), Sun (el viento, la mayor de las hijas) se opone a Chen (el trueno, el hijo mayor), y Li (el sol, la hija segunda) se opone a K'an (la luna, el hijo segundo). Si interpretamos los trigramas en un sistema binario, partiendo del que ocupa el centro de la parte superior, el Ch'ien, y en sentido contrario a las agujas del reloj, tenemos la secuencia numeral: 0, 1, 2 y 3; partiendo de abajo, del K'un, y subiendo en sentido contrario a las agujas del reloj, tenemos la secuencia: 7, 6, 5 y 4.

Suscripciones:

Prensa Científica, S.A.
Departamento de suscripciones
Calabria, 235-239
Barcelona-15 (España)

Distribución para España:
Distribuciones de Enlace, S.A.
Ausias March, 49
Barcelona-10 (España)

Distribución para los restantes países:
Editorial Labor, S.A.
Calabria, 235-239
Barcelona-15 (España)

Condiciones de suscripción:

España:
Un año (12 números): 1.540 pesetas
Seis meses (6 números): 800 pesetas

Extranjero. Un año (12 números): 30 U.S. $

Ejemplar atrasado ordinario:
160 pesetas
Ejemplar atrasado extraordinario:
225 pesetas.

PROCEDENCIA DE LAS ILUSTRACIONES

Portada: Ilustración original de Ilil Arbel

Página	Fuente	Página	Fuente
6-11	Gabor Kiss	80	James S. Albus y John M. Evans, Jr., del National Bureau of Standards
15-18	Gabor Kiss	81	Jon Brenneis
21-30	Allen Beechel	82-83	Ralph Morse
32	David E. Comings y Tadashi Okada, City of Hope Medical Center	84	Dan Todd
34-35	George V. Kelvin	86	James S. Albus y John M. Evans, Jr., del National Bureau of Standards
36	Gary S. Stein, Facultad de Medicina de la Universidad de Florida	87	Ralph Morse
37-43	George V. Kelvin	89	Carol Donner
45	Antonio Romañá	90-91	Total Communications Industries, Ltd.
50-53	Ramón Margalef	92-95	Carol Donner
55-58	Ramón Margalef	96-99	Total Communications Industries, Ltd.
59	Antonio Vidal	100	William A. Garnett
61	Juan Armengol	102-107	Eric G. Hieber Associates, Inc.
64-69	Tom Prentiss	108	*Progrès Agricole et Viticole*
70	Cía. Mexicana Aerofoto	110	Asunción López
71-74	Tom Prentiss	111-116	Ilil Arbel
76	Ralph Morse	123	Humberto Fernández-Morán Universidad de Chicago
78	Auto-Place, Inc.		
79	Dan Todd		

Dep. legal: B. 38.999-76
Fotocomposición: Fort, S.A.
Rosellón, 33 - Barcelona-15
Cayfosa. Santa Perpetua de Moguda. Barcelona.

Printed in Spain - Impreso en España

PRESENTACION

Porque ciencia es pensamiento contrastado, creación intelectual sometida a la continua decantación provocada por hechos nuevos y por nuevas interpretaciones de los ya conocidos, comunicación y ciencia son inseparables. Hechos y teorías adquieren su pleno valor de saber objetivo cuando pueden ser comprobados, comprendidos y aceptados por cualquier persona racional, al menos en el caso ideal. Frente a todo tipo de saberes esotéricos, ciencia es conocimiento no sólo compartible, sino compartido de hecho. Por ello también, ciencia y secreto son términos antagónicos.

En una época como la nuestra, en que el esfuerzo investigador se realiza con una intensidad y una abundancia de recursos personales y materiales casi inimaginables en épocas pasadas, la difusión de sus resultados tiene una importancia primordial. Por un lado, como parte esencial del propio proceso científico. Nunca ha estado más patente que ahora el carácter colectivo de la tarea científica, la necesidad que tiene el innovador de basarse en y tener en cuenta los resultados obtenidos por otras personas que trabajan su misma especialidad e incluso otras, afines o muy alejadas de ella. Y nunca tampoco ha sido más difícil estar bien informado de lo que se ha hecho y se está haciendo, paradójicamente por exceso de información, por la misma riqueza y abundancia de los datos que continuamente se dan a conocer a través de innumerables publicaciones periódicas, libros, congresos o conferencias en todo el mundo.

Pero hay otro aspecto de la comunicación científica no menos importante y, sin embargo, muy reciente. A lo largo del siglo XIX *y de modo muy especial en el presente, la ciencia, a través de la tecnología, ha ido influenciando más y más todos los aspectos de la vida del hombre, ya se trate de su salud, su alimentación, su trabajo, sus medios de transporte o de información, sus diversiones, la organización de sus sistemas sociales o económicos, o la idea que tiene de su ubicación en este planeta y en el cosmos. La ciencia ha adquirido así un papel social de primera magnitud y, en consecuencia, sus interrelaciones con la sociedad son múltiples y complejas. La sociedad exige de la comunidad científica que dé respuesta a los interrogantes teóricos existentes y contribuya a resolver los problemas prácticos que están planteados. Como contrapartida, la sociedad tiene que establecer las condiciones que hagan posible el trabajo científico. Una fluida comunicación entre la comunidad científica y la sociedad en general es requisito indispensable para que esto pueda lograrse.*

Si el intercambio de información entre los científicos ya plantea serios problemas, el que debe haber entre ellos y la sociedad es mucho más arduo todavía; incluso es de otra naturaleza. A la necesidad de selección cuantitativa y cualitativa se une ahora la de cambiar de lenguaje, la de pasar del obligado tecnicismo en que se elabora y expresa la ciencia al sistema de referencias del no especialista. Hay que traducir con la mayor fidelidad posible las

expresiones científicas a otras accesibles al hombre culto, cosa que es factible aunque no sencilla. Este es precisamente el objetivo que INVESTIGACION Y CIENCIA *perseguirá desde ahora incansable en cada uno de sus números.*

Para conseguirlo, no escatimará esfuerzos. Cuenta, por una parte, con la colaboración de científicos de todo el mundo, cuya voz autorizada expondrá con la máxima sencillez y rigor los resultados de sus investigaciones más recientes. Y cuenta también con el interés de sus lectores, dispuestos a realizar el esfuerzo exigido por un discurso riguroso a cambio de compartir el patrimonio en constante aumento del conocimiento humano. Cumplirá su misión si logra ser el intermediario eficaz entre unos y otros.

Ni España ni los países latinoamericanos pueden vanagloriarse de una brillante tradición científica. Las causas son sociales e institucionales, que no genéticas o geográficas. Cualquier persona trabajando en cualquier región del mundo puede hacer progresar la ciencia. Es más, hay que reconocer el mérito de lo conseguido por nuestros investigadores a pesar de desarrollar su labor en ambientes tan poco favorables. Y sabido es que el nivel científico de muchos españoles, argentinos, mexicanos o brasileños que trabajan en instituciones de otros países en nada desmerece del de sus colegas extranjeros. La falta de canales de comunicación científica en nuestro ámbito cultural es notoria y grave, porque no sólo es efecto sino también causa de esta anemia científica y, por tanto, cultural.

INVESTIGACION Y CIENCIA *nace con la ilusión de ser un factor del cambio que necesariamente debe producirse. Gracias a la colaboración entusiasta y decidida de quienes tan ejemplarmente hacen* SCIENTIFIC AMERICAN, *ofrecerá mes a mes los resultados más sobresalientes de la investigación mundial. Incluida, por supuesto, la que con tanto empeño y tan poco reconocimiento social realizan los investigadores de nuestros propios países. La ciencia y el saber no conocen fronteras y en parte alguna se encuentra tan desplazado como en ellos un mal entendido orgullo nacionalista. Es mucho lo que podemos y debemos aprender de lo que se está haciendo en otros países. Y mucho también lo que ignoramos de lo que se hace en los nuestros.* INVESTIGACION Y CIENCIA, *que se dirige primordialmente al lector culto, especializado o no, se esforzará por facilitar al máximo este proceso de comunicación en la seguridad de que así no sólo favorece el desarrollo de la ciencia, sino también, y muy especialmente, el progreso social, que no es factible si los ciudadanos libres y con participación en los asuntos colectivos no disponen de la suficiente información para guiar su vida y la de los grupos a que pertenecen, teniendo en cuenta el atributo más específico y más valioso de la humanidad: la inteligencia transformada en cultura.*

FRANCISCO GRACIA

Los autores

J. RICHARD GOTT III, JAMES E. GUNN, DAVID N. SCHRAMM y BEATRICE M. TINSLEY ("¿Expansión indefinida del universo?") colaboran juntos desde 1974, año en que Gott y Gunn enseñaban en el Instituto de Tecnología de California y Schramm y Tinsley en la Universidad de Texas, en Austin. Gott se graduó en 1969 en la Universidad de Harvard con la calificación de *summa cum laude* y obtuvo en 1972 su Ph. D. en astrofísica por la Universidad de Princeton; y recientemente se ha incorporado al claustro de profesores de Princeton. Gunn, graduado en la Universidad de Rice, recibió su Ph. D. en el Cal Tech, en 1965. Enseña astronomía en el Cal Tech y es miembro del consejo asesor de los observatorios Hale. Schramm se doctoró por el Cal Tech en 1971. Actualmente es profesor asociado de astronomía y astrofísica en el Instituto Enrico Fermi de la Universidad de Chicago. Tinsley adquirió su MS. en física en la Universidad de Canterbury, Nueva Zelanda (1963) y su Ph. D. en astronomía por la de Texas (1967). Desempeña en la actualidad el cargo de profesora asociada de astronomía en la Universidad de Yale.

DAVID B. CLINE, ALFRED K. MANN y CARLO RUBBIA ("La búsqueda de nuevas familias de partículas elementales") prosiguen su colaboración a pesar de la geografía. Cline enseña física en la Universidad de Wisconsin, donde obtuvo su Ph. D. en 1965. Mann profesa en la Universidad de Pennsylvania, a la que se trasladó en 1949, dos años después de obtener su Ph. D. en la de Virginia. Rubbia ejerce la docencia en la Harvard University; de origen italiano, completó su doctorado en la Universidad de Pisa y trabajó durante ocho años en el Centro Europeo para la Investigación Nuclear (CERN) antes de pasar a ser miembro del claustro docente de Harvard en 1970.

GARY S. STEIN, JANET SWINEHART STEIN y LEWIS J. KLEINSMITH ("Las proteínas cromosómicas y regulación de los genes") han trabajado en equipo durante varios años a pesar de su distante separación geográfica. Gary Stein recibió su grado de Doctor por la Universidad de Vermont en 1969, y pasó tres años en el departamento de patología de la Facultad de Medicina de la Temple University, antes de trasladarse a Florida en 1972. Janet Swinehart Stein se licenció en química por el Elizabethtown College de Pensilvania en 1968 y trabaja en el doctorado en la Princeton University. Kleinsmith, profesor de zoología, en la Universidad de Michigan, obtuvo su grado de Doctor por la Universidad Rockefeller en 1968, después de pasar dos años en la facultad de medicina.

RAMON MARGALEF LOPEZ ("Biología de los embalses") es, en la actualidad, catedrático de Ecología y Limnología en la Universidad de Barcelona, en cuya ciudad nació en 1919. Durante toda su vida ha trabajado en numerosos proyectos de investigación, siendo director del Instituto de Investigaciones Pesqueras a lo largo del bienio 1966-1967. Merecen destacarse de su amplia bibliografía: *Los organismos indicadores en limnología* (1955), *Perspectives in Ecological Theory* (1968) y *Ecología* (1974). Ha aplicado la teoría de la información al estudio de las comunidades naturales y ha formulado modelos matemáticos propios para la dinámica de una población. Margalef reconoce la ayuda que le han prestado, en el trabajo de campo que le ha servido de base para este artículo, sus colaboradores en el Departamento de Ecología de la Universidad de Barcelona, dentro de un programa de investigación patrocinado por la Dirección General de Obras Hidráulicas.

JEREMY A. SABLOFF y WILLIAM L. RATHJE ("El surgimiento de una clase mercantil maya") son, respectivamente, profesor agregado de Antropología en la Universidad de Harvard y profesor agregado de Antropología en la Universidad de Arizona. Sabloff, que desempeña también el cargo de conservador adjunto de arqueología mesoamericana en el Harrards Peabody Museum of Archeas' and Ethnology, se graduó en la Universidad de Pennsylvania y obtuvo el doctorado en Harvard. Rathje obtuvo su título de licenciado en 1967 en la Universidad de Arizona y su Doctorado en 1971 en Harvard. Entre otras actividades, es Director de un proyecto de la Universidad de Arizona que investiga los desperdicios de los hogares modernos para ver si la información obtenida de una civilización conocida puede sacarse también de una civilización antigua, aplicando técnicas similares.

JAMES S. ALBUS y JOHN M. EVANS, JR. ("Sistemas robot") trabajan en el Office of Developmental Automation and Control Tecnology, del National Bureau of Standards. Albus estuvo 15 años en la National Aeronautics and Space Administration (NASA), para pasar luego al National Bureau de Standards; estudió física en el Wheaton College y, una vez graduado, ingeniería eléctrica en las Universidades de Ohio y Maryland. Ha escrito varios trabajos teóricos sobre los mecanismos neurológicos del cerebro. Evans, que entró en el Bureau de Standards en 1970, se graduó en la Universidad de Yale y está en posesión del Ph. D. en la especialidad de Física por la Universidad de Colorado.

JOHN CAIRNS ("El problema del cáncer") trabaja en los Mill Hill Laboratories del Imperial Cancer Research Fund en Londres. Tras graduarse en medicina por la Universidad de Oxford, se dedicó a la investigación, durante varios años, sobre la multiplicación de los virus y la visualización del ADN por autorradiografía. Más tarde pasó a los Estados Unidos como director del Cold Spring Harbor Laboratory de Biología cuantitativa hasta 1968, en que James D. Watson le sustituyó. Durante cierto tiempo continuó trabajando en la síntesis del ADN.

PHILIP WAGNER ("Vinos, vides y clima") es director jubilado del *Sun*, de Baltimore; sigue ejerciendo su oficio de periodista en un diario sindical en donde escribe un comentario político, sobre gestión gubernamental, dos veces por semana. "La viticultura ha constituido para mí una segunda vocación durante largos años. Es una ocupación ideal en total disparidad con mi profesión y ambas ocupaciones las mantengo sin ningún tipo de interferencias". Su trabajo en viticultura, en estrecha colaboración con su esposa, se ha centrado en el cultivo de viñas, experimentos con nuevas variedades, multiplicación de vides y elaboración de vinos.

¿Expansión indefinida del universo?

La recesión de las galaxias lejanas, la densidad media de la materia, la edad de los elementos y la cantidad de deuterio muestran que la expansión es continua e irreversible

J. Richard Gott III, James E. Gunn, David N. Schramm y Beatrice M. Tinsley

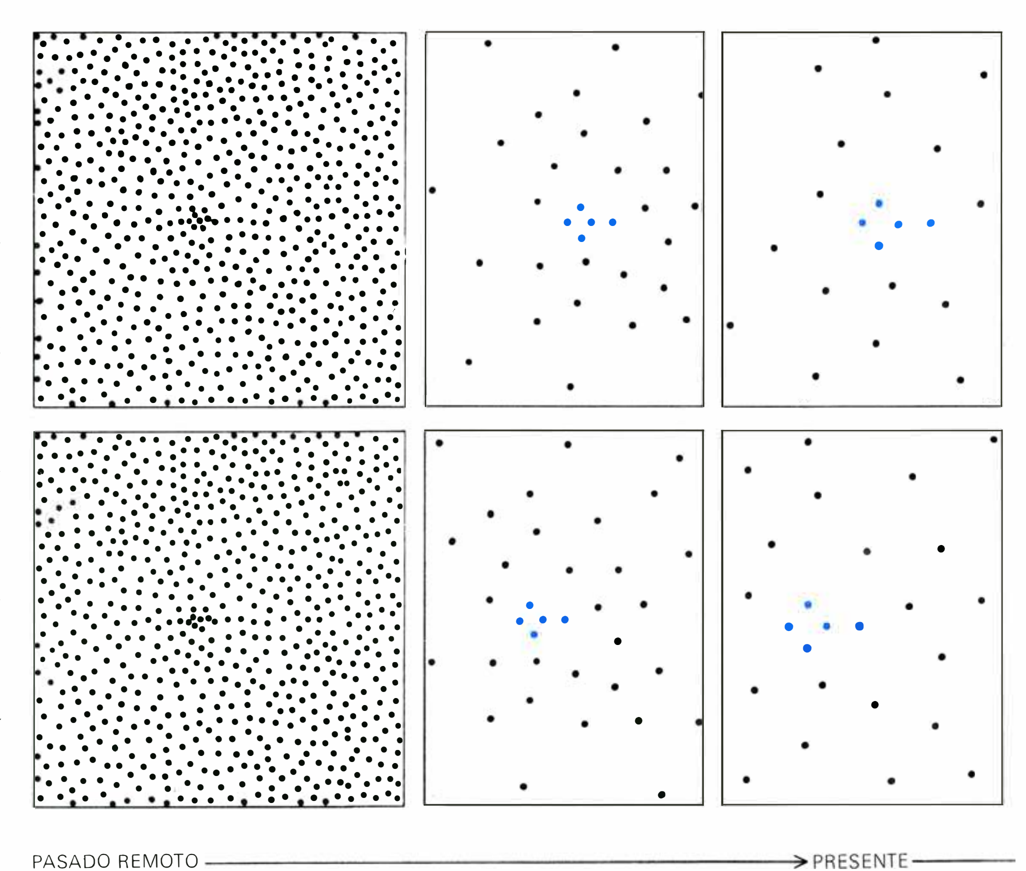
La investigación cosmológica es muy antigua, pero fue en los últimos 50 años cuando se inició el conocimiento real de cómo empezó el universo y cuál pueda ser su destino futuro. La observación crucial ocurrió por el año 1920, en que Edwin P. Hubble demostró que las nebulosas espirales no eran objetos locales sino sistemas de estrellas independientes, muy parecidas a las nuestras; de ello dedujo que el universo era mucho mayor que lo creído hasta entonces. Más tarde Hubble probó que el conjunto de galaxias observables se movía ordenadamente. Hoy ya se sabe con certeza que este movimiento es de expansión: todas las galaxias lejanas se van distanciando de nosotros.

Se considera probado, pues, que el universo está en expansión. Un problema pendiente de resolución es saber si la expansión continuará siempre o si las galaxias que se están alejando detendrán, algún día, su proceso e invertirán luego la dirección de su movimiento para caer en un gran colapso. La respuesta al problema implica la determinación del carácter geométrico del universo, es decir, condiciona la naturaleza del espacio y del tiempo. Si la expansión continuara indefinidamente, el universo sería "abierto" e infinito; pero si se detuviera e invirtiera la dirección, el universo sería "cerrado" y de extensión finita.

Para optar por una u otra posibilidad, los astrónomos construyen modelos matemáticos del universo, con la esperanza de que la observación de los hechos del universo real confirme su modelo y excluya los demás. Todavía no se ha logrado que un solo tipo de medida resuelva el problema con precisión suficiente y sin ambigüedad. Se pueden aducir, sin embargo, varias pruebas independientes, y se han encontrado elementos del rompecabezas por parte de muchos observadores que emplean técnicas completamente distintas. Sería buena cosa que reuniéramos hoy esas piezas dispersas. Agrupadas ordenadamente, manifiestan una evidencia plausible: el universo es abierto y nunca cesará en su expansión.

Expansión isótropa

Hubble llegó al descubrimiento del movimiento de recesión de las galaxias lejanas a través de las mediciones de sus espectros ópticos. Los espectros de la mayoría de las estrellas (y por tanto de las galaxias) están cortados por líneas oscuras, que corresponden a la absorción de ciertas longitudes de onda por parte de los átomos situados en capas más frías y exteriores de la atmósfera estelar; cada elemento químico produce un espectro de líneas característico cuyas longitudes de onda pueden medirse con precisión en el laboratorio. Cuando la galaxia se aleja del observador,



EL ESTADO FUTURO DEL UNIVERSO se describe mediante modelos matemáticos de su modo de evolución. Se consideran aceptables, en general, dos clases de modelos; en ambos, el universo se origina en estado compacto de densidad infinita (la gran explosión). Para una clase de modelos, el universo se expande continua e independientemente con un ritmo de velocidad decreciente *(serie su-*

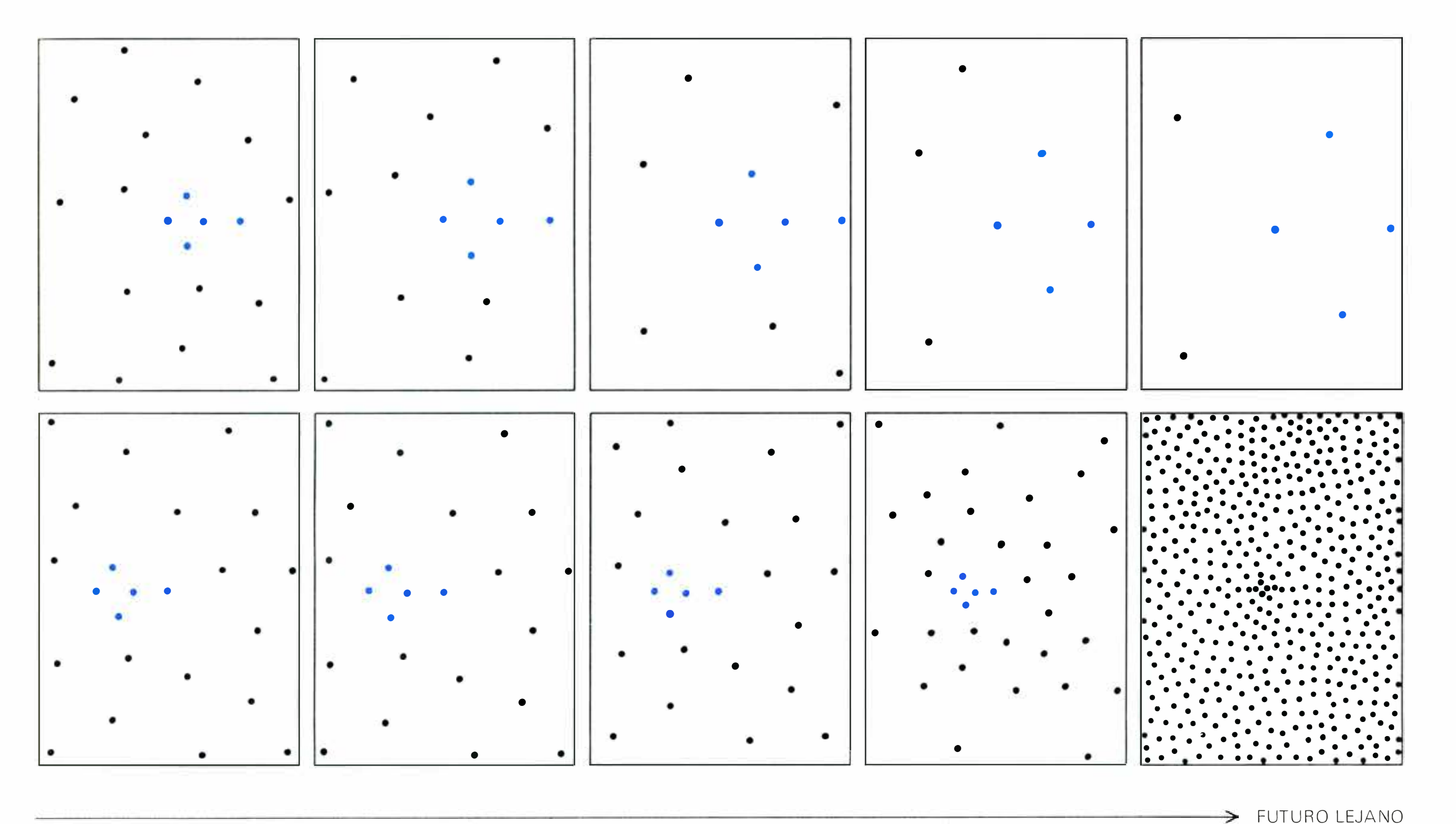
la longitud de onda de cada línea del espectro aumenta en virtud del efecto Doppler, de modo que todas las líneas aparecen desplazadas hacia longitudes de onda más largas y en particular hacia el extremo rojo de la zona visible del espectro. A este desplazamiento se le llama corrimiento hacia el rojo, y de la medida de su magnitud podemos deducir la velocidad de recesión. En virtud del efecto Doppler, cuando un objeto se mueve hacia el observador, las longitudes de onda de las líneas espectrales decrecen; las líneas aparecen desplazadas hacia el extremo azul del espectro, efecto que recibe el nombre de corrimiento hacia el azul. Todos los espectros de las galaxias lejanas que fueron medidos por Hubble y luego por otros observadores dan corrimientos hacia el rojo; de lo cual se deduce, pues, que las galaxias estarían alejándose de nosotros.

El movimiento de recesión tiene varias propiedades interesantes. Hubble demostró que la velocidad con que una galaxia se aleja es proporcional a su distancia de nosotros, de modo que puede calcularse el cociente constante entre distancia y velocidad. La razón es de forma tal que una galaxia distante 10 millones de años-luz de nosotros se aleja con una velocidad de 170 kilómetros por segundo; otra galaxia doblemente distante se alejará a doble velocidad, o sea, a 340 kilómetros por segundo *(véanse las ilustraciones inferiores)*. Se observan frecuentemente pequeñas desviaciones a esta regla, ya que la mayoría de las galaxias forman parte de grupos o cúmulos y tienen movimientos orbitales con una componente sobre la línea de visión que une la Tierra con la galaxia. Con todo, estos movimientos se distribuyen al azar, de suerte que, si tomamos un conjunto grande de galaxias, éstas se anulan mutuamente. Se han encontrado variaciones sistemáticas del cociente, no debidas por tanto al azar, únicamente para las galaxias situadas a distancias extremas; como veremos, estas variaciones no sólo no invalidan la regla sino que aportan una información importante para la historia del universo.

Una segunda característica de la expansión cósmica es su isotropía: se mantiene igual en todas las direcciones. Sin que influya su situación en el firmamento, la velocidad de recesión de una galaxia está relacionada con su distancia, en virtud de la misma ley de proporcionalidad. De esta observación podría desprenderse que el universo posee una simetría muy notable y, lo que aún es más extraordinario, que nosotros estamos en el mismo centro.

Hay otra explicación, por supuesto. La comprenderemos rápidamente si nos fijamos en un modelo sencillo, de dos dimensiones, de un universo en expansión. Imaginemos un globo esférico con puntos pintados en su superficie; cada punto representa una galaxia. Al hinchar el globo, aumenta la distancia entre dos puntos cualesquiera (tomamos siempre las medidas en la superficie de la esfera) con una velocidad proporcional a la distancia que hay entre ellos. Sea cual sea el punto que designemos como centro, todos los demás puntos se van alejando del mismo, uniformemente, en todas las direcciones. Una expansión tal no tiene ningún centro; o, con mayor precisión, todos los puntos son centros.

De este análisis de la expansión se deduce que no puede cambiar la configuración geométrica de los puntos. Un globo en el que se hubiera pintado a Mickey Mouse, continuaría con el mismo dibujo una vez hinchado. Las distancias entre los distintos puntos del globo se multiplican por el mismo factor. Algo parecido ocurre en el universo real; si en un



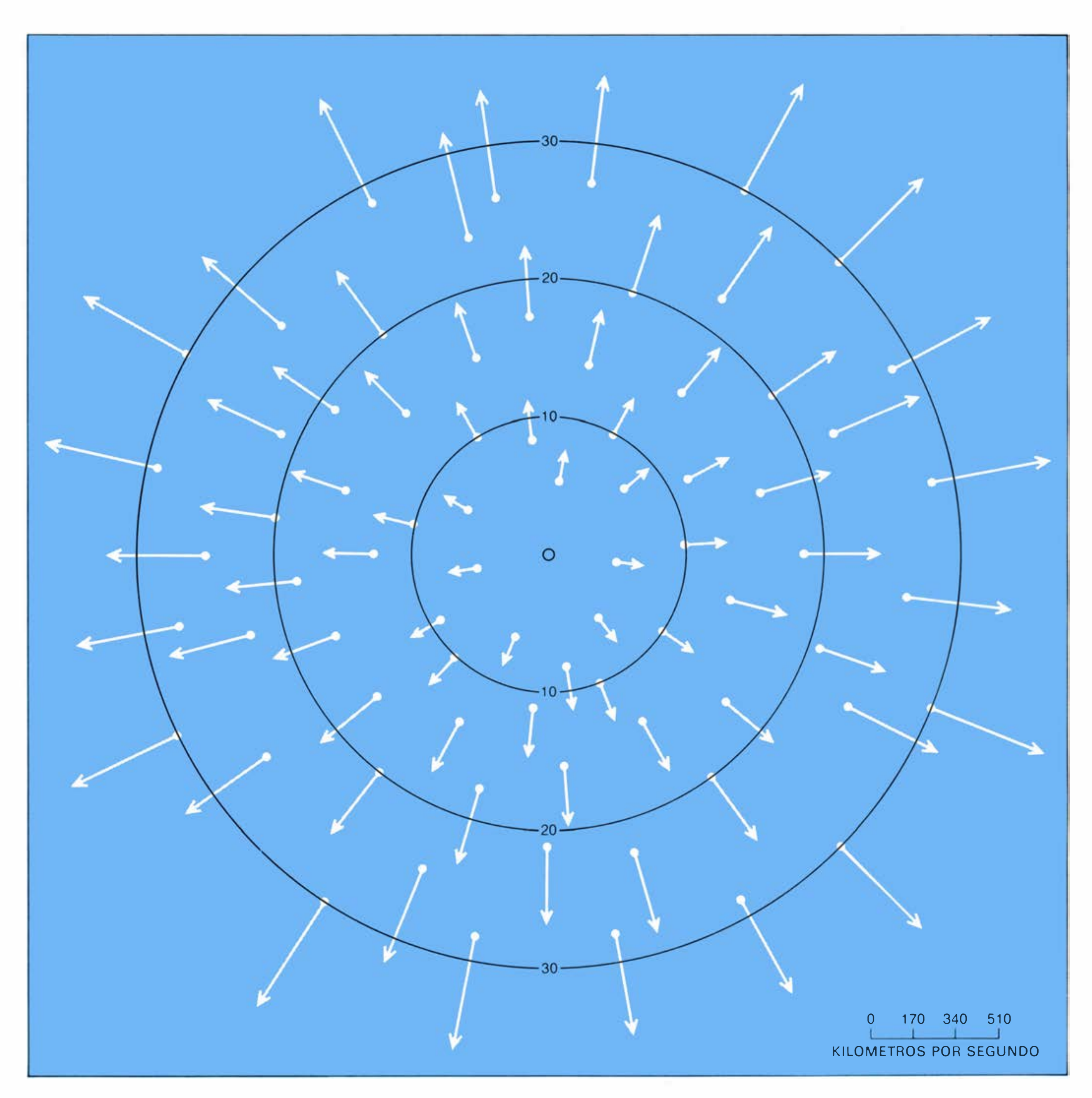
perior); para la otra, se expande hasta alcanzar un volumen máximo; empieza luego a contraerse volviendo a una potencial densidad infinita *(serie inferior de dibujos)*. Se representa aquí una y otra alternativa, tomando una región arbitraria del espacio en donde la expansión (isótropa) cósmica se simboliza por la disminución de la densidad media. Así pues, un observador situado en cualquier punto se encontrará como centro de la expansión; asimismo la forma de cualquier conjunto (el conjunto arbitrario de puntos coloreados) será la misma en cualquier época.

instante determinado ocho galaxias se encuentran en los vértices de un cubo, permanecerán en los vértices de un cubo cada vez mayor, conforme el universo vaya expansionándose.

La gran explosión *(Big Bang)*

A partir del descubrimiento primero de Hubble, observaciones cada vez más precisas han demostrado que no sólo la expansión es isotrópica: todos los aspectos a gran escala del universo son indiferentes a la dirección. Por ejemplo, la distribución de las galaxias sobre la esfera celeste y la distribución de las fuentes de radio que no son de nuestra galaxia se muestran completamente uniformes. El caso más evidente de isotropía fue descubierto en 1965 por Arno A. Penzias y Robert W. Wilson, de los laboratorios Bell; se trata de la radiación de fondo de microondas que, al parecer, impregna a todo el universo. Se ha demostrado que la radiación de microondas tiene una gran isotropía, pues consideradas todas las áreas del firmamento, varía en menos de una sobre mil partes.

La observación de tan notable isotropía ha inducido a aceptar una generalización muy potente conocida por principio cosmológico. Este principio afirma que el universo se muestra isotrópico para todos los observadores que participan en la expansión, sin que importe lugar ni tiempo en que se hallen. Dicho



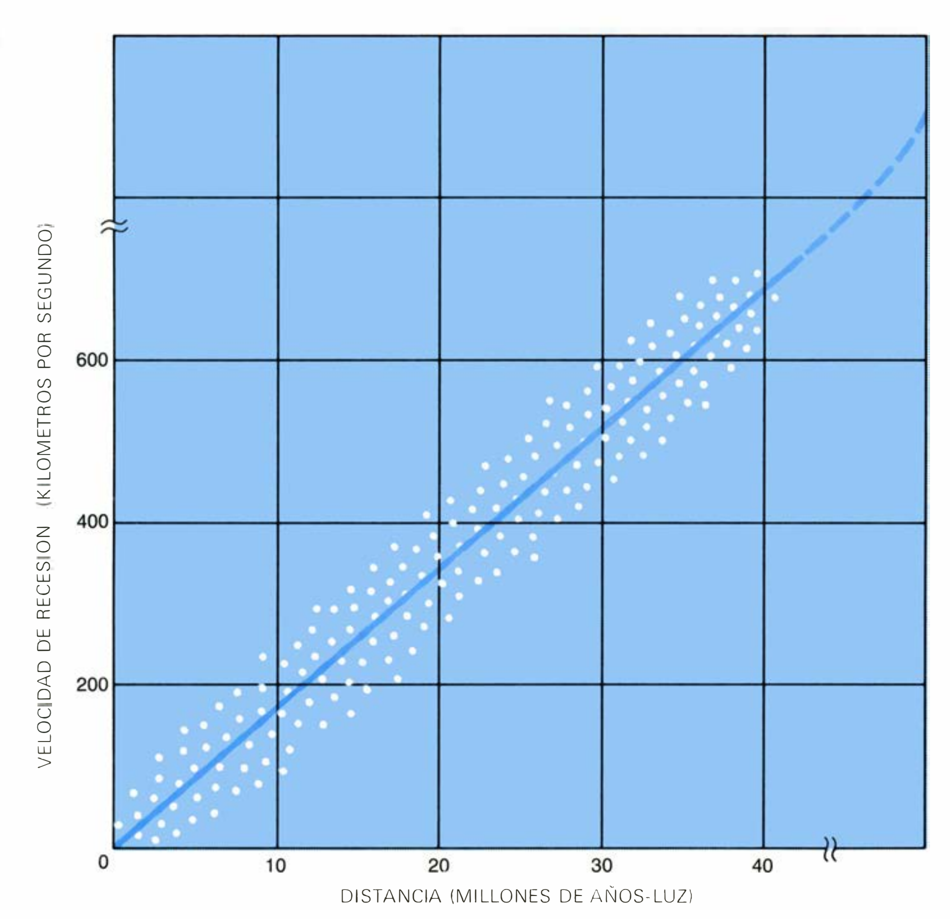
LA EXPANSION COSMICA parece colocar al observador en el centro del universo, del que las galaxias distantes se van alejando. La velocidad con que una galaxia se aleja es proporcional a su distancia al observador, relación que estableció por primera vez, en 1920, Edwin P. Hubble a partir de observaciones hechas con el telescopio de 254 cm del Mount Wilson Observatory. Desde entonces se conoce por ley de Hubble el principio según el cual el cociente entre la velocidad y la distancia es constante. Puede interpretarse llanamente como prueba de que la expansión del universo empezó con la gran explosión ya que la relación implica que en el pasado las galaxias formaban un todo compacto con una densidad infinita. Las distancias se expresan en millones de años luz y las flechas dan el valor de las velocidades (escala del extremo inferior).

de otra manera, nuestra galaxia ocupa el centro del universo, pero eso mismo puede decirse de cualquier otra galaxia.

El principio cosmológico regula también la conducta del modelo de universo, de dos dimensiones, representado por el globo esférico. Si hemos distribuido, con densidad uniforme, los puntos pintados en la superficie del globo, el entorno de cualquier punto que escojamos es estadísticamente igual al de otro punto cualquiera, sin que ninguna dirección tenga un significado especial. No tendría sentido postular, de un modo independiente, que los puntos (o, en el universo tridimensional, las galaxias) estén disdistribuidos uniformemente; si el universo es isótropo para todos los observadores, la distribución debe ser homogénea; si no fuera así, el observador que estuviera situado en el límite de una región donde hubiera una fluctuación de la densidad no vería una distribución uniforme independiente de la dirección.

Para los propósitos de nuestra investigación aceptaremos el principio cosmológico, recordando empero que su mordiente es de orden preferentemente filosófico. Ni se ha probado contundentemente ni tal vez puedan darse jamás pruebas concluyentes del mismo.

Con los datos que hoy nos ofrecen las observaciones del universo, ¿qué podemos decir de su historia? Empecemos adoptando un modelo hipotético, sencillo, que consista en suponer que la velocidad de recesión de cada galaxia fue siempre la misma. De acuerdo con ello, la galaxia que ahora está lejos de nosotros estuvo alguna vez muy cerca, y que el tiempo transcurrido desde entonces es igual al cociente de la distancia de la galaxia por su velocidad. Puesto que dicho cociente es el mismo para todas las galaxias, resulta que todas debieron haber estado juntas en un cierto instante; en otras palabras, en un instante dado del pasado, toda la materia del universo estuvo de tal forma comprimida que poseía una densidad prácticamente infinita. Se llama *tiempo de Hubble* al tiempo que se estima ha transcurrido desde ese instante de estado compacto, suponiendo que la intensidad de la expansión no ha variado. Se llama *constante de Hubble* al recíproco del tiempo de Hubble; multiplicando la constante de Hubble por la distancia a la que está una galaxia obtenemos su velocidad de recesión. Las medidas del tiempo de Hubble son difíciles debido a las imprecisiones en el cálculo de las distancias de las galaxias; se han revisado periódicamente desde la primera estimación de Hubble, que fue de unos dos mil millones de años. Hoy



LA LEY DE HUBBLE se deduce midiendo el cociente entre la velocidad y la distancia de muchas galaxias. El mejor cálculo de este cociente *(línea coloreada continua)* **da el valor de unos 17 kilómetros por segundo por millón de años luz. Las galaxias, tomadas una a una** *(puntos blancos),* **no se ajustan a ese valor porque la mayoría forman parte de cúmulos y tienen velocidades orbitales. El inverso de la razón es el tiempo de Hubble: el tiempo que tardaría cualquier galaxia, a la velocidad actual, en llegar a su posición actual, es decir, el tiempo desde la gran explosión si las velocidades no hubiesen cambiado. Existe la creencia de que las velocidades de recesión han disminuido debido a la gravitación; por tanto, el cociente crecerá para grandes distancias** *(línea coloreada a trazos).*

se cree que el tiempo de Hubble está entre doce y veinticinco mil millones de años, siendo el valor más aceptado el de diecinueve mil millones de años.

Si extrapolamos el movimiento de las galaxias hacia el pasado más remoto posible, llegaríamos a un estado en el que las galaxias estarían comprimidas con una densidad infinita. Este estado corresponde a la gran explosión, y señala el origen del universo y de todo lo que hay en él.

Con cálculos sencillos de proporcionalidad, si las velocidades de recesión no hubiesen cambiado, deducimos que la gran explosión debió ocurrir hace exactamente un tiempo de Hubble. Es casi seguro que la velocidad de recesión no sea constante en realidad, pero ello no implica la no existencia de la gran explosión, simplemente cambia la fecha del mismo.

Si partimos de que las leyes físicas que conocemos son válidas y, en algún sentido, completas, síguese como conclusión inevitable que el mundo comenzó con una gran explosión. Pudiera darse que la naturaleza tuviera leyes cuyos efectos resultaran despreciables en un laboratorio de física o incluso en el sistema solar, pero fueran predominantes a la hora de determinar el comportamiento del universo como un todo. Un conjunto doctrinal que requería nuevas leyes de la naturaleza era la cosmología del estado estacionario según la cual el universo nunca cambia y es infinito en el tiempo. Para explicar la expansión cósmica, la teoría del estado estacionario postulaba la creación continua de materia a partir del vacío.

En el modelo de estado estacionario del universo resultaba muy difícil explicar la radiación de fondo de microondas. Este campo de radiación tiene unas características espectrales semejantes a las de la radiación térmica emitida por un cuerpo que estuviera a la temperatura de 2,7 grados Kelvin. La única explicación satisfactoria, al parecer, sería interpretarlo como residuo de una época en que el universo se hallara a temperatura y densidad muy elevadas. Ahora bien, ningún modelo de estado estacio-

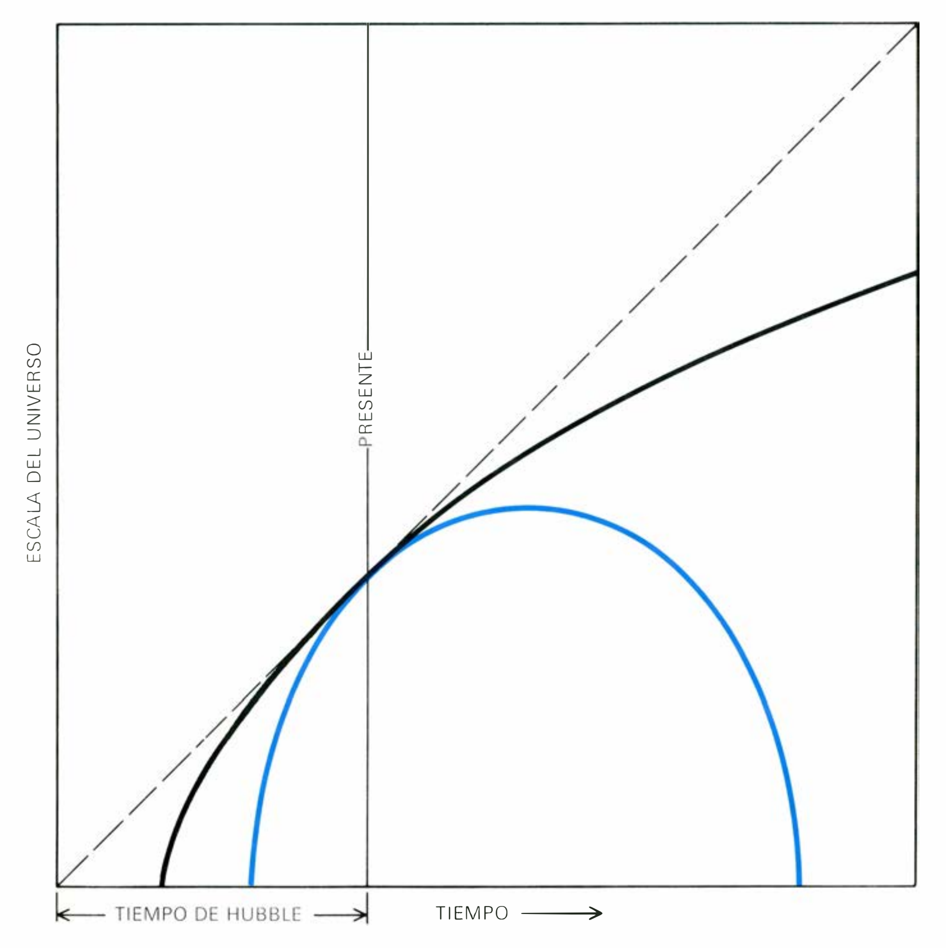
nario pudo atravesar esas condiciones, pues en dicho modelo, por definición, nunca cambian las condiciones.

En los modelos de la gran explosión la radiación de fondo es consecuencia natural de las condiciones existentes en el universo primitivo. Al estado inicial de dichos modelos le corresponde una gran densidad y una gran temperatura; ese estado recibe a veces el nombre de granada *(fireball)* cósmica. Se cree que, en esa fase, la materia y la energía electromagnética que formaban el universo estaban en equilibrio termodinámico y, por tanto, el espectro de la radiación era el de un cuerpo sometido a altas temperaturas. Al ir expansionándose el universo, la radiación se fue enfriando, dando lugar al espectro de baja temperatura que observamos hoy. Este enfriamiento puede interpretarse como un enorme corrimiento hacia el rojo; puesto que las galaxias van distanciándose de la radiación, su espectro se está desplazando constantemente hacia longitudes de onda más largas, correspondientes a energías y temperaturas más bajas. En 1946, George Gamow predijo la existencia de una radiación térmica de fondo basándose únicamente en el marco teórico del modelo de la gran explosión. Sus cálculos daban una temperatura actual de esta radiación de unos cinco grados Kelvin. La concordancia, en líneas generales, entre la predicción de Gamow y las observaciones de Penzias y Wilson constituye la prueba más convincente de la gran explosión. Parece, pues, que el universo se desarrolló a partir de un estado de densidad infinita, hará de ello sobre un tiempo de Hubble aproximadamente. El espacio y el tiempo se crearon entonces, así como toda la materia del universo. La pregunta inmediata de qué ocurrió antes de la gran explosión no tiene sentido; sería como preguntarse qué hay al norte del Polo Norte. En el mismo orden, sería insensato preguntar dónde ocurrió la gran explosión. El "universo-punto" no era algo aislado en el espacio; encerraba a todo el universo. La única respuesta posible, por tanto, habría de ser que la gran explosión ocurrió por todas partes.

Si escogemos para su análisis una región esférica del universo, podemos considerar al resto del universo que la envuelve como una capa esférica hueca, ya que el principio cosmológico exige que la materia envolvente esté distribuida uniformemente en todas direcciones. Podemos estudiar, pues, la esfera seleccionada como si se encontrara aislada y no estuviera sometida a ninguna fuerza exterior. Además, el principio cosmológico sostiene que, sea cual sea la esfera de galaxias que seleccionemos, debe expansionarse o contraerse según el mismo factor con que lo hace todo el universo, sin que importe el lugar en que se halle ni el tamaño de la esfera. A la hora de abordar la dinámica del universo, sólo necesitaremos, consiguientemente, observar la dinámica de una esfera representativa que esté dentro de él. Si la esfera escogida es pequeña, las velocidades de las galaxias serán muy inferiores a la velocidad de la luz, y podremos describir sus movimientos por medio de la mecánica de Newton.

Una galaxia situada en el límite de esta pequeña región esférica estaría sometida únicamente a las fuerzas gravitatorias generadas por la materia interior a la esfera. Si esta materia está distribuida homogéneamente, la fuerza resultante que actúa sobre la galaxia debe atraerla hacia el centro de la esfera. En consecuencia, la galaxia escogida no se aleja a velocidad constante sino que su movimiento de recesión es siempre decelerado. Resulta inmediato, por tanto, que, en el pasado, la galaxia escogida y todas las demás restantes debieron desplazarse con mayor rapidez que actualmente. Prescindir de la deceleración conduce a una sobreestimación de la edad del universo. Esta edad sería un tiempo de Hubble si la velocidad de expansión no hubiera cambiado; ahora bien, bajo la influencia de la gravitación la velocidad ha disminuido, por consiguiente la gran explosión hubo de ocurrir en un tiempo inferior a un tiempo de Hubble.

En la mayoría de modelos de un universo en evolución, se supone que las galaxias que se van distanciando siguen trayectorias balísticas parecidas a las que describe el lanzamiento de una pelota o una granada. Las galaxias se propulsaron por fuerzas que actuaron en el momento de la gran explosión, y desde entonces se mueven en caída libre sin ninguna otra propulsión. Deberían con-



LOS MODELOS DE EVOLUCION COSMICA describen los cambios en la escala del universo con el paso del tiempo. Todos los modelos deben estar de acuerdo con la escala y velocidad de expansión que hoy se observa, de modo que sus respectivas gráficas han de cortar el momento presente. Cuando la velocidad de expansión no cambia *(línea negra a trazos),* **la edad del universo es el tiempo de Hubble. Los universos que pierden aceleración son más jóvenes, y tanto su historia como su futuro dependen de la magnitud de la deceleración. Si la deceleración es pequeña, la expansión puede continuar indefinidamente, aunque con velocidad decreciente** *(línea negra continua).* **Una deceleración mayor implica que la expansión cósmica se detenga e invierta luego su sentido conduciendo a un posible colapso** *(línea de color).* **Se llama "abierto" al universo que se expande indefinidamente y, "cerrado", al universo que está abocado al colapso, que es también el más joven.**

	ABIERTO	CRITICO	CERRADO
PARAMETRO DE DENSIDAD Ω $\frac{\text{DENSIDAD ACTUAL}}{\text{DENSIDAD CRITICA}}$	$\Omega < 1$	$\Omega = 1$	$\Omega > 1$
PARAMETRO DE DECELERACION q_0 DECELERACION $\frac{\text{DISTANCIA}}{(\text{VELOCIDAD})}$	$q_0 < \frac{1}{2}$	$q_0 = \frac{1}{2}$	$q_0 > \frac{1}{2}$
GEOMETRIA DEL ESPACIO	HIPERBOLICA (CURVATURA NEGATIVA)	PLANA (CURVATURA NULA)	ESFERICA (CURVATURA POSITIVA)
FUTURO DEL UNIVERSO	EXPANSION PERPETUA	EXPANSION PERPETUA	COLAPSO POSIBLE

LOS MODELOS ABIERTOS Y CERRADOS del universo se distinguen principalmente por la densidad media de materia y por el valor de la deceleración cósmica. La densidad es crucial ya que en los modelos descritos por la teoría de la relatividad general es la única causa determinante de las fuerzas gravitatorias que frenan la expansión cósmica. Esta magnitud se comprende mejor usando un parámetro sin dimensiones, que se obtiene del cociente de la densidad actual y la densidad crítica, donde esta última es la exigida para detener la expansión. La deceleración puede expresarse también por un número, el parámetro de deceleración, que en los modelos descritos vale la mitad del parámetro de densidad. Estos dos parámetros determinan el futuro del universo y la geometría del espacio. El universo abierto es siempre infinito, su espacio tiene curvatura hiperbólica, negativa. En el universo con densidad crítica, donde el parámetro de densidad vale uno, el espacio tiene curvatura cero, y es el espacio plano con geometría euclídea. El universo cerrado es finito; su espacio tiene curvatura esférica o positiva.

tinuar, pues, en movimiento uniforme siempre que no existan otras fuerzas que actúen sobre ellas. Si sólo permitimos que aparezcan, en nuestros modelos, las fuerzas normales que dan lugar a las leyes físicas conocidas, la única fuerza que puede causar un efecto significativo en la expansión es la gravitación. Por consiguiente, comprenderemos la dinámica del universo en expansión cuando logremos describir las interacciones gravitatorias de todas sus componentes.

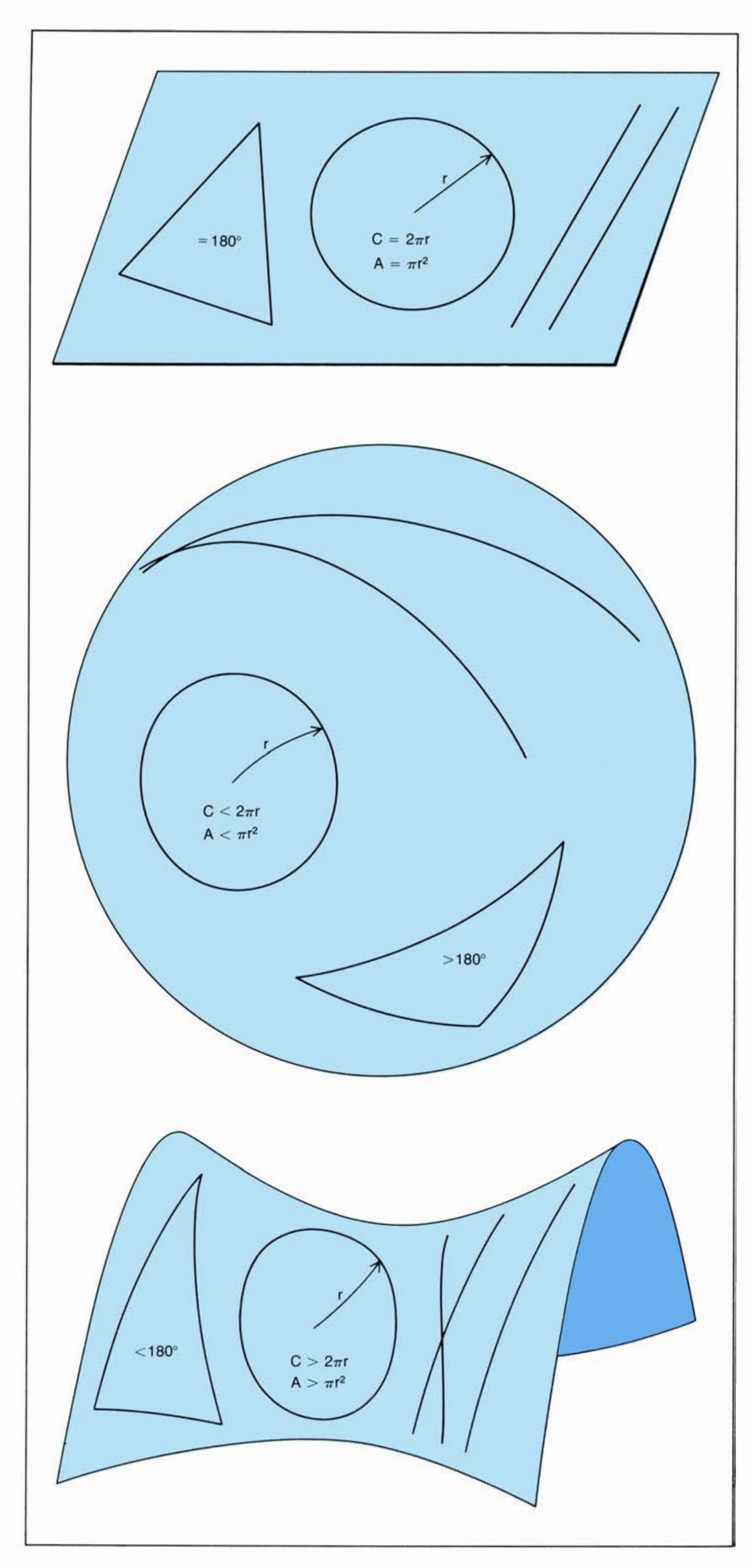
La deceleración gravitatoria

Las fuerzas gravitatorias actúan sobre toda la materia, son siempre positivas y su alcance es infinito. La gravitación tiene, además, una propiedad geométrica peculiar que facilita grandemente su estudio, a saber: la esfera hueca no ejerce ninguna fuerza resultante de tipo gravitatorio sobre las masas que se introduzcan en ella. (La masa de la capa esférica atrae, por supuesto, a las masas del interior, pero todas las fuerzas se anulan mutuamente, de modo que en cualquier punto del interior la fuerza resultante es cero.) Newton fue quien primero demostró dicha propiedad, si bien puede aplicarse a las teorías de la gravitación más recientes, verbigracia, la teoría de la relatividad general.

El valor de la deceleración gravitatoria depende de la cantidad de masa que encierre la esfera seleccionada. Si la esfera contiene mucha materia, la galaxia escogida llegará a detenerse y caer hacia el centro; la región esférica afectada empieza a colapsarse, y, aplicando el principio cosmológico, igual proceso seguirá el universo entero. Pero si hubiera poca materia, la galaxia escogida sufrirá una deceleración continua sin llegar nunca a detenerse; tanto la región esférica como el universo en su totalidad se expansionarán indefinidamente. La situación es análoga a la producida por el disparo que se va distanciando de la superficie terrestre: el proyectil decelera pero no volverá a la superficie si su velocidad inicial es superior a un cierto valor crítico, la velocidad de escape.

La velocidad de escape para los objetos que se alejan de la tierra depende de la masa de ésta; para una galaxia escogida en el límite de una esfera arbitraria del espacio, la velocidad de escape depende de la masa total que hay dentro de la esfera. Del cociente entre la velocidad y la distancia deducimos la velocidad actual de la galaxia escogida respecto del centro de la esfera. Su suerte final depende, pues, del valor de la velocidad de escape y, por tanto, de la masa encerrada en la esfera.

Como hemos supuesto que el universo es homogéneo, la cantidad decisiva será la densidad media de la materia del universo. Si la densidad fuera menor que un cierto valor crítico, la fuerza de gravitación sería demasiado pequeña para detener la expansión cósmica, y todas las galaxias se alejarían para siempre (aunque cada vez con mayor lentitud. (Si la densidad fuera mayor que la densidad crítica, la fuerza de gravitación prevalecerá y se frenará la expansión hasta pararse, comenzando entonces una contracción acelerada que terminaría en un cataclismo final al que podríamos llamar el gran colapso *(big crunch)*. El valor actual de la densidad crítica depende del tiempo de Hubble, que no conocemos



con precisión. Si el tiempo de Hubble vale diecinueve mil millones de años, la densidad crítica será de 5×10^{-30} gramos por centímetro cúbico, que corresponde a unos tres átomos de hidrógeno por metro cúbico. Esta densidad parece pequeñísima, pero debe recordarse que en promedio el universo está muy vacío.

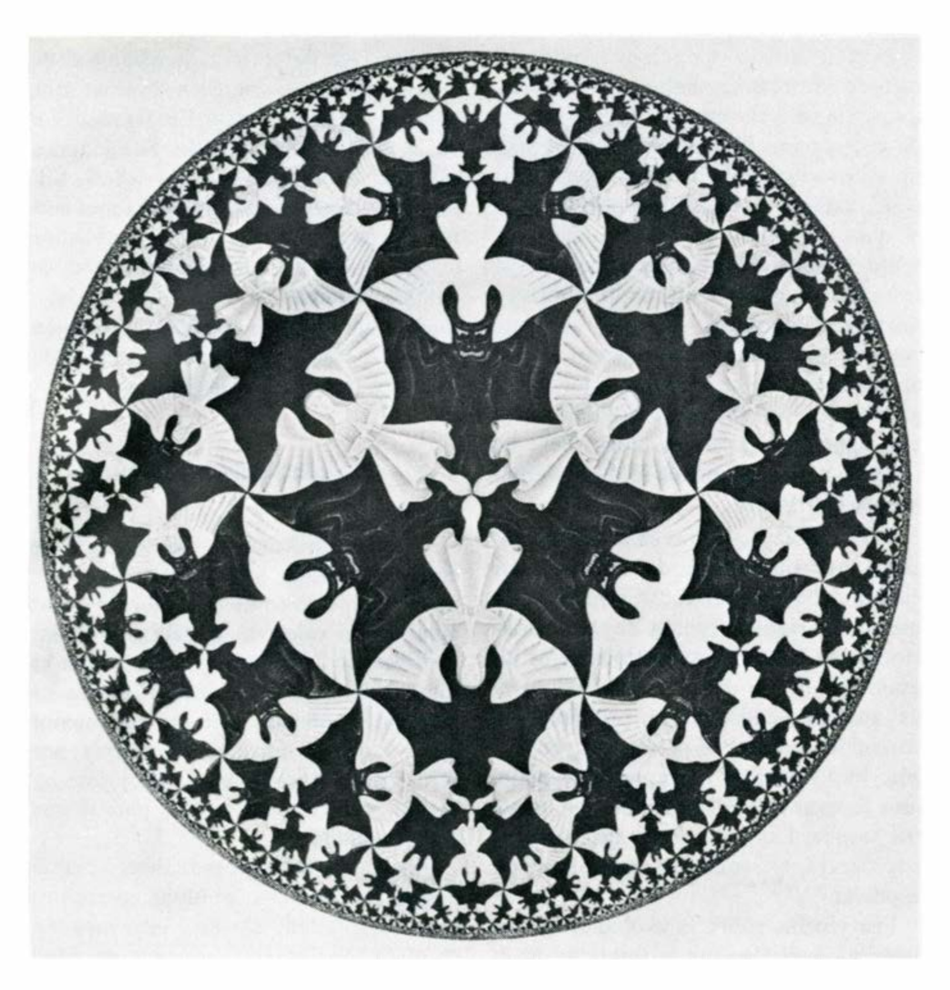
Si utilizamos un número sin dimensiones por parámetro de densidad, al que designaremos con la letra griega omega (Ω), podemos introducir el efecto de la gravitación sobre la expansión cósmica en modelos matemáticos. El parámetro de densidad define la razón entre la densidad actual del universo y la densidad crítica. Si el universo se expansionara indefinidamente, esta razón habrá de ser menor o igual a uno; si Ω fuera igual a uno, el universo se iría expansionando a la misma velocidad de escape, y, si Ω es mayor que uno, el universo tendería irremisiblemente al colapso.

La geometría del espacio

El razonamiento anterior podría haberse deducido en su totalidad de la teoría de la gravitación de Newton, aunque también es válido en la teoría de la relatividad general. Si bien en ésta el valor del parámetro densidad tiene consecuencias adicionales; en particular, determina la geometría del espacio. En el universo de densidad elevada, condenado al colapso, la gravitación es suficientemente fuerte para "cerrar" el espacio. El volumen total del universo en cualquier instante es finito, por más que no haya ningún límite o borde del mismo. Un análogo en dos dimensiones de esta superficie tridimensional es la superficie de la esfera, que, parecidamente, posee un área finita sin tener ningún borde.

Cuando Ω es igual a uno, es decir, cuando el universo se expande a la misma velocidad de escape, la geometría del espacio es "plana". Nos hallamos en la conocida geometría de Euclides, y, en dos dimensiones, el análogo sería un plano

LA GEOMETRIA DEL ESPACIO asociada a cada modelo de universo guarda una analogía con las superficies. Las propiedades de las superficies derivan de los axiomas y teoremas de Euclides sobre las líneas paralelas, sobre la suma de los ángulos de un triángulo y sobre el perímetro de una circunferencia y el área de un círculo. El espacio plano correspondiente al universo crítico se representa por un plano y el espacio con curvatura positiva de un universo cerrado se corresponde con la superficie de una esfera. Algunas propiedades de un espacio con curvatura negativa de un universo abierto pueden compararse con las de la superficie "silla de montar", pero ésta es una analogía imperfecta porque tiene un centro. La mejor representación de un universo abierto sería una superficie infinita, llamada pseudoesfera, que no puede construirse en un espacio tridimensional.



SUPERFICIE DE UNA PSEUDOESFERA, según el grabado *Círculo límite IV* de M.C. Escher. En él, la superficie está proyectada sobre un plano. Por tratarse de un mapa obtenido por proyección, la escala no es constante; en la pseudoesfera real todas las figuras de ángeles y demonios tendrían el mismo tamaño. Si tomamos una figura como unidad de medida resulta obvio que la longitud de la circunferencia crece proporcionalmente más rápida que su radio. Análogamente, cada figura define un triángulo (con vértices en los pies y en las puntas de las alas); a partir del número de triángulos que concurren en un vértice puede demostrarse que, en una pseudoesfera, la suma de los ángulos de un triángulo es menor que 180 grados. La pseudoesfera es una superficie infinita de curvatura negativa, análoga al espacio de un universo en expansión indefinida. No tiene ninguna posición privilegiada que sea el centro y la proyección sería la misma de haberla enfocado sobre otro punto.

infinito. Es mucho más difícil representar la geometría de un universo en expansión perpetua en que Ω sea menor que uno. Su análogo en dos dimensiones sería la superficie de una figura que se ha dado en llamar pseudoesfera; pero no se puede construir una pseudoesfera completa en un espacio tridimensional. Algunas propiedades de este espacio las tendría una superficie en forma de silla de montar, pero el modelo no valdría porque esta superficie tendría un centro, mientras que el espacio real no posee ninguna posición privilegiada *(véase la ilustración de la pág. 12)*. Tal vez la mejor representación en dos dimensiones de este espacio sea la proyección de una pseudoesfera sobre un plano, procedimiento que empleó en varias de sus obras el artista holandés M. C. Escher *(véase la ilustración de esta misma página)*.

Las tres clases posibles de espacios tridimensionales se distinguen por sus propiedades geométricas, algunas de las cuales podemos representar en modelos análogos de dos dimensiones. El plano llano constituye el fundamento de la geometría euclidiana y en él se cumplen todos los axiomas de Euclides y los teoremas que de ellos se derivan. Por un punto dado de un plano sólo puede trazarse una recta que sea paralela a otra recta; los ángulos de un triángulo suman siempre 180 grados; la longitud de una circunferencia aumenta proporcionalmente al radio y el área de un círculo aumenta proporcionalmente al cuadrado del radio.

En la superficie de una esfera no hay rectas paralelas, siempre que definamos como línea recta aquella que nos da el camino más corto entre dos puntos. Estas líneas reciben el nombre de geodésicas, y, en la esfera son los círculos máximos, dándose siempre una intersección entre dos cualesquiera de ellos. De igual modo: en una esfera la suma de los ángulos de un triángulo es siempre mayor que 180 grados; la longitud de una circunferencia aumenta, en proporción, más lentamente que el radio; y el área de un círculo aumenta, en proporción, más lentamente que el cuadrado del radio.

La superficie de una pseudoesfera posee propiedades opuestas a las de una esfera. Por un punto dado pueden trazarse infinitas rectas que sean paralelas a otra recta o geodésica. La suma de los ángulos de un triángulo es menor que 180 grados. La longitud de una circunferencia aumenta, en proporción, más rápidamente que el radio, y el área de un círculo aumenta, en proporción, más rápidamente que el cuadrado del radio. Nikolai Lobachevski fue el primero en estudiar, en 1826, la geometría de un espacio tridimensional representada por una pseudoesfera.

En los modelos cosmológicos sencillos aquí revisados la geometría del espacio se halla estrechamente relacionada con la evolución futura del universo. Conviene destacar que, en los modelos con Ω mayor que uno el universo es cerrado en espacio y en tiempo. El volumen del espacio es finito; para el tiempo, existen unos límites muy precisos, a saber: empieza con la gran explosión y acaba con el gran colapso. Los modelos en los que Ω es menor o igual a uno son abiertos en espacio y en tiempo. Estos modelos parten de un punto muy preciso, la gran explosión, pero tienen siempre una extensión infinita y se expandirán indefinidamente en el futuro.

Mediciones de la deceleración

Hay varios caminos posibles para determinar si el universo actual es abierto o cerrado. Todos ellos conducen, en definitiva, a una estimación de la intensidad con que la expansión cósmica está decelerando. Un método sencillo consiste en medir directamente la deceleración fijándose en las galaxias lejanas. Otro estribaría en calcular la edad del universo y deducir la variación de la velocidad de expansión al comparar dicha edad con el tiempo de Hubble (que sería la del universo de no existir deceleración). Por ser la deceleración un fenómeno gravitatorio, tendríamos una medida equivalente en la densidad media de la materia; la comparación de la densidad actual con la densidad crítica da el cociente Ω. Por último, la abundancia actual de ciertos elementos químicos es una suerte de recuerdo fósil de las condiciones existentes en el universo primitivo,

densidad incluida, de cuya información podemos también deducir el valor de Ω. Nuestro conocimiento actual del estado del universo se funda en los resultados obtenidos con todos y cada uno de estos métodos.

La deceleración de la expansión cósmica suele expresarse en función de un número sin dimensiones llamado parámetro de deceleración, representado por q_o. Como la deceleración es un efecto gravitatorio, el parámetro de deceleración está estrechamente relacionado con la densidad media de la materia. En los modelos cosmológicos que aquí consideramos, todos ellos deducidos de acuerdo con la teoría de la relatividad general, q_o es siempre exactamente igual a la mitad del parámetro de densidad Ω. De modo que, si q_o es mayor que $\frac{1}{2}$, y debido a su elevada densidad, el universo iría reduciendo su velocidad hasta detener la expansión para caer en el colapso. Si q_o es menor que $\frac{1}{2}$, la expansión no puede pararse porque la densidad sería demasiado pequeña para hacerlo.

Un procedimiento obvio para determinar la deceleración sería medir dos veces la velocidad radial de una galaxia cualquiera y observar cuánto ha disminuido en el intervalo. Desgraciadamente, el cambio de velocidad esperado durante el tiempo de una vida humana es mínimo y no puede detectarse; y, sin duda, los errores experimentales anexos a la medición son varios órdenes de magnitud mayores que el valor esperado. Con todo, por ser finita la velocidad de la luz, se pueden calcular las velocidades que tenían las galaxias en un pasado remoto y compararlas con sus velocidades en épocas más recientes. Esta comparación es factible, pues al observar en el firmamento objetos cada vez más lejanos los estamos situando también en un pasado cada vez más remoto. Tal razonamiento es inmediato cuando medimos las distancias en años-luz: si una galaxia se halla a una distancia de mil millones de años-luz, la luz que de ella recibimos hoy se emitió hace mil millones de años; en su espectro por el efecto Doppler, debe reflejarse la relación entre la velocidad de la galaxia entonces y nuestra velocidad ahora. Por consiguiente, si la expansión cósmica está decelerando, no hay que esperar que el cociente entre velocidad y distancia, constante según descubriera Hubble, siga siendo tal para las galaxias más lejanas. Cuando se trate de grandes distancias, el cociente debe ir aumentando, es decir, las velocidades habrán de ser mayores que las deducidas de la ley de Hubble.

Para medir la deceleración por este método es necesario haber calculado antes, independientemente, las distancias de las galaxias. El único método práctico de estimación de la distancia para todas ellas, excepto para las más próximas, es el que se basa en su luminosidad aparente. Si todas las galaxias, en cualquier instante, tuvieran la misma luminosidad intrínseca, entonces su brillo aparente variaría simplemente en razón inversa al cuadrado de su distancia; por consiguiente, el cálculo de la distancia sería una operación sin dificultades. Pero no todas las galaxias tienen la misma luminosidad intrínseca.

Variaciones aleatorias en el brillo (causadas por diferencias de tamaño, por ejemplo) pueden provocar errores en la medición. Habida cuenta de tales variaciones, se hace necesario considerar un gran número de datos, sometiéndolos a un análisis estadístico; en principio, las variaciones aleatorias no constituyen ningún impedimento grave, ya que podemos esperar que se anulen en una muestra amplia. Las variaciones sistemáticas, por contra, sí requieren una corrección explícita.

Las teorías sobre la evolución de las estrellas sugieren que la totalidad de la luz procedente de las estrellas de una galaxia aislada debe disminuir, probablemente, según una proporción de pocas unidades sobre ciento, cada mil millones de años. Resulta verosímil, pues, que, en un pasado lejano, las galaxias fueran más brillantes. Si al medir la deceleración se despreciara este cambio en el brillo, las distancias calculadas serían excesivamente pequeñas, sobreestimándose en consecuencia el valor de la deceleración. Parece que la disminución del brillo sea mínima, pero en realidad cambia, en más o menos 1, el valor calculado del parámetro de deceleración q_o, lo cual es más que suficiente para decidirse por un universo abierto o cerrado. Las observaciones actuales más fidedignas, que tienen en cuenta el cambio de luminosidad de las estrellas con su evolución, sostienen que q_o se halla más cerca de cero que de $\frac{1}{2}$ y, por tanto, que el universo es abierto y en expansión indefinida.

Hay otra dificultad de peso en el cálculo de la deceleración. La mayoría de las galaxias observadas están en cúmulos relativamente densos; por ello, deberían tenerse en cuenta las interacciones entre galaxias. Por ejemplo, recientemente se ha demostrado que, en los cúmulos, las grandes galaxias absorben a las más pequeñas, resultando un cambio en la luminosidad y en el tamaño. No es posible precisar por ahora la magnitud de este cambio, ni siquiera afirmar que comporte un incremento o disminución en la luminosidad medida. Si añadimos estrellas a una galaxia, ésta debería brillar más, pero en las observaciones cosmológicas se mide sólo la luminosidad de la parte central de la galaxia; cuando la galaxia "caníbal" trague o absorba bastante, el número de estrellas en la región central tenderá a reducirse, despidiendo entonces menos brillo la galaxia.

La edad del universo

Como consecuencia del margen de error estadístico y de nuestro conocimiento imperfecto de la evolución de las galaxias, el valor de q_o deducido de las medidas de la velocidad de recesión es muy incierto. De esta única prueba no podemos concluir que q_o sea menor que $\frac{1}{2}$ y que el universo sea abierto; por otro lado, valores más elevados para q_o, verbigracia q_o igual a dos, parece que deben excluirse.

El segundo enfoque para determinar el destino del universo se funda en el cómputo de su edad. De no existir ninguna deceleración de la expansión, su edad sería de un tiempo de Hubble. Pero, como hay deceleración, debe ser algo menor que un tiempo de Hubble. Por consiguiente, si hallamos la diferencia entre la edad real y el tiempo de Hubble se podrá calcular en principio el parámetro de deceleración q_o.

Hay dos métodos para estimar la edad del universo; si bien sólo dan límites inferiores, porque únicamente calculan las edades de objetos del universo y estos objetos se formaron, probablemente en los primeros mil millones de años a partir de la gran explosión. Con el primer método se trata de determinar la edad de las estrellas más viejas observables todavía. Se cree que las estrellas más viejas, y suficientemente próximas como para poder observarse con detalle, son las que están en cúmulos globulares asociados a nuestra propia galaxia. Los modelos de evolución estelar indican que su edad debe oscilar entre ocho y dieciséis mil millones de años.

La edad puede estimarse también a partir de la cantidad relativa de ciertos elementos pesados. Se supone que todos los elementos más pesados que el hierro incluyendo algunos radiactivos, se formaron en las supernovas; las cuales, a su vez, se hallan, presumiblemente, explotando dentro de la galaxia desde su formación. Al desintegrarse cada elemento radiactivo con una intensidad cons-

tante, el cociente entre la cantidad de cada elemento y la cantidad de sus productos de desintegración puede darnos la edad media de los elementos pesados. Los valores deducidos muestran que la edad de la galaxia debe estar entre seis y veinte mil millones de años. Las dos edades calculadas son, pues, compatibles, y ambas sugieren que la gran explosión hubo de ocurrir en un tiempo oscilante entre los ocho y los dieciocho mil millones de años.

La densidad media

El que una determinada edad, dentro del margen permitido, corresponda a un universo abierto o cerrado depende del valor del tiempo de Hubble, valor que, según se ha visto, no es fácil de determinar. Aun cuando tomemos como valor del tiempo de Hubble la mejor estimación reciente que lo cifra en diecinueve mil millones de años, ni los límites de la edad ni la exclusión de valores de q_0 mayores que dos son suficientes para decidir si el universo es abierto o cerrado *(véase la ilustración de la pág. 16)*. Para sacar una conclusión necesitamos que se den, además, otras condiciones.

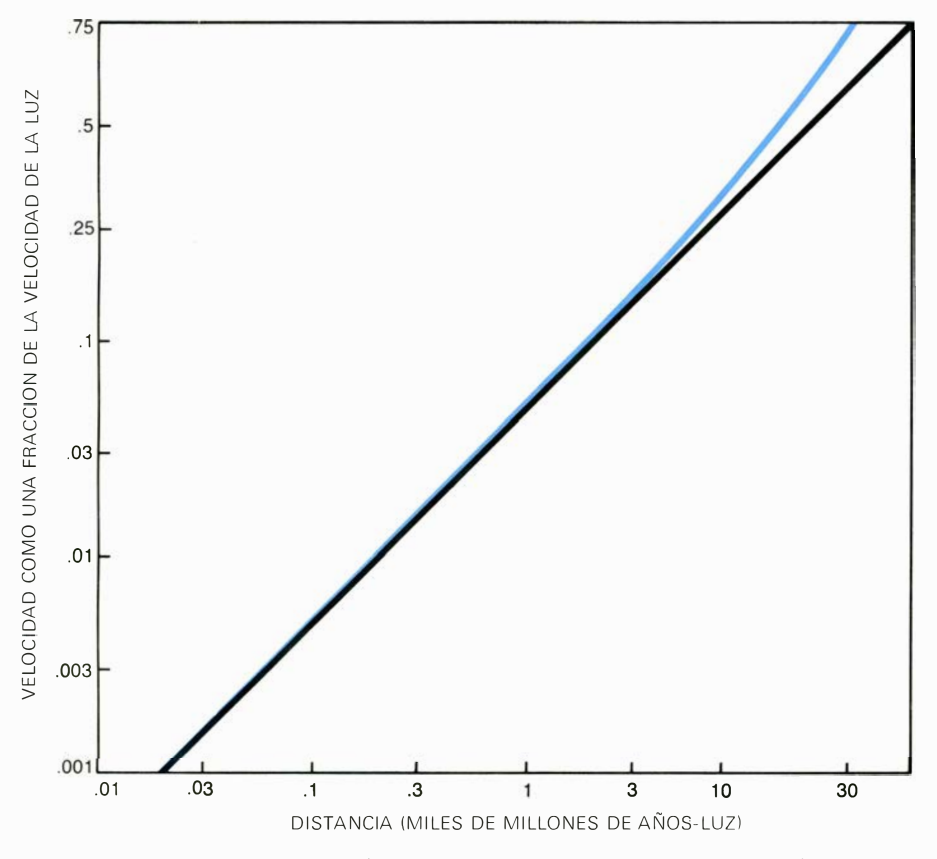
La tercera prueba consiste en medir la densidad media de la materia del universo, y de ahí deducir el parámetro de densidad Ω. Podemos obtener un límite inferior para la densidad considerando sólo la masa asociada con las galaxias visibles. Esta densidad se halla multiplicando el número de galaxias que hay en un volumen dado del espacio por la masa de las galaxias y dividiendo por el volumen.

No es tan difícil pesar una galaxia como parece a primera vista. Existen muy pocas galaxias totalmente aisladas; la mayoría están en pequeños grupos o en grandes cúmulos, y podemos deducir sus masas de las observaciones de las interacciones gravitatorias que se ejercen mutuamente. Sean, por ejemplo, dos galaxias que giran una alrededor de la otra: entre ellas debe haber una atracción gravitatoria que compense la fuerza centrífuga. Si conocemos su distancia, y la velocidad de cada una respecto de la otra, la determinación de la suma de sus masas resulta un ejercicio muy simple de mecánica de Newton. El procedimiento a seguir en el caso de un cúmulo de muchas galaxias apenas encierra mayor complicación. Conviene resaltar que la masa así calculada no comprende sólo la masa de las galaxias sino también la masa total de cualquier otra materia que se halle en el cúmulo. Las partes que no sean visibles, tales como agujeros negros o polvo y gas extragaláctico, quedan automáticamente incluidas.

El cálculo de la masa de un gran número de galaxias, junto con el recuento del número de ellas en grandes volúmenes del espacio, da una estima del valor del parámetro de densidad Ω. Si la masa total del universo fuera la referida a las galaxias, el valor de Ω rondaría el 0,04, y el universo sería claramente abierto y en estado de expansión indefinida. Este valor podría modificarse en un factor tres, de modo que un valor de Ω equivalente a 0,12 sería todavía compatible con las observaciones, valor que sigue siendo mucho menor que 1, escalar requerido para cerrar el universo.

La densidad del universo puede deducirse también de la comparación del comportamiento de las galaxias lejanas con el comportamiento de las que se encuentran en el supercúmulo local, que es el agregado de galaxias que comprende nuestro propio grupo local amén de otros muchos grupos pequeños y el cúmulo algo mayor de Virgo. En el supercúmulo local, la densidad media de galaxias es dos veces y media mayor que la de todo el universo. Si toda la masa perteneciera a las galaxias, la densidad media de la materia debería ser también dos veces y media mayor en el supercúmulo que fuera de él. Esta diferencia en el valor de la densidad causaría aquella otra diferencia en la intensidad de expansión, pues si la densidad local es mayor, las galaxias que estén próximas a otras irán amortiguando su aceleración más rápidamente. La magnitud de esta diferencia depende del valor de Ω; si Ω es grande, la diferencia será considerable. Ahora bien, si Ω es pequeña, la deceleración será pequeña en cualquier parte, e incluso en zonas donde la densidad aumenta en un factor dos y medio el cambio sería pequeño. En realidad la diferencia no puede detectarse al ser menor que los errores probables de observación. La conclusión más evidente es que Ω es muy pequeño, a buen seguro no mucho mayor que 0,1.

Ambos métodos de cálculo de la densidad limitan su objetivo a la materia asociada con las galaxias; a los cuales, se podría objetar que existen plausiblemente cantidades sustanciales de materia en otras partes del universo. Sin embargo, aunque se pueda excluir esa posibilidad, no se tiene ningún argumento que lo pruebe.

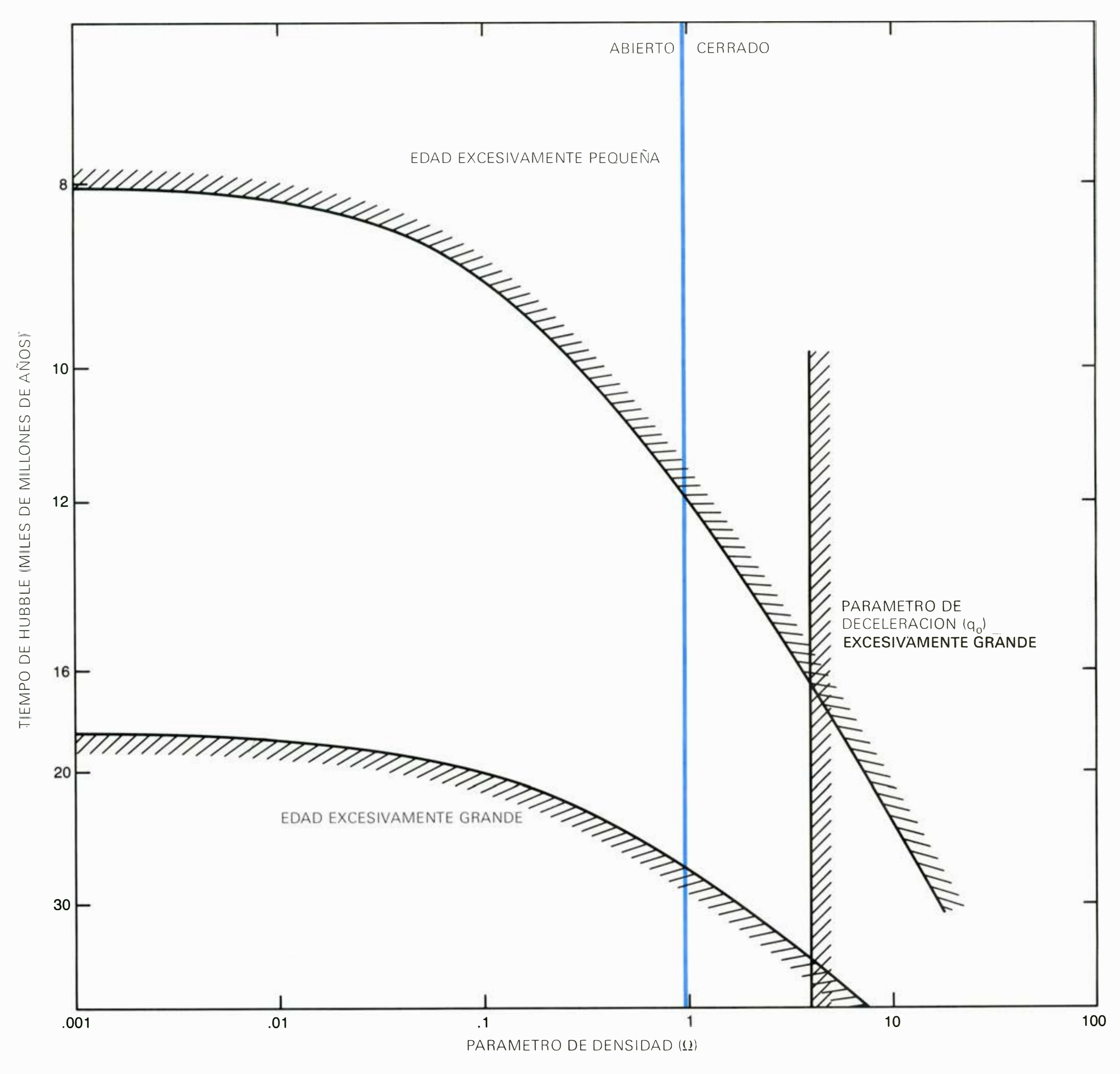


LA DECELERACION de la expansión cósmica puede detectarse gracias a las velocidades de recesión de las galaxias en un pasado remoto. Se puede inspeccionar el pasado observando las galaxias más lejanas, ya que la luz que ahora nos llega se emitió hace mucho tiempo, determinado por la distancia de la galaxia en años luz. La deceleración se contempla pues como una desviación de la ley de Hubble: si aquélla no existiera, la razón de la velocidad respecto a la distancia sería constante *(línea negra)*; con deceleración, la razón crece para distancias extremas *(línea de color)*. Por ser muy difícil medir la distancia de las galaxias es imposible obtener con precisión el valor de la deceleración, pero se han excluido los valores del parámetro de deceleración mayores que 2.

Las teorías actuales demuestran que los cúmulos de galaxias podrían haberse formado a partir de un universo en el que la materia estuviera distribuida mucho más uniformemente que tal como está ahora. Si en la formación de galaxias se hubieran producido residuos, también éstos hubiesen sido atrapados por los cúmulos. Así pues, toda partícula que hoy no se halle incluida en los cúmulos tiene que ser porque los evitó ante todo; es decir, para escapar a la captura, las partículas hubieron de tener una posición y una velocidad iniciales que fuesen especiales y fuera de lo normal. Si ahora hubiera una gran cantidad de materia distribuida uniformemente fuera de los cúmulos, caería en ellos en los miles de millones de años inmediatos.

Otra vía alternativa sería pensar que la masa necesaria estuviese formada por algún medio distribuido uniformemente, el cual se hallara a una presión interna capaz de contrarrestar la fuerza gravitatoria de las galaxias. Podría tratarse por ejemplo, de un gran número de neutrinos o de ondas gravitatorias. Con todo, existe un argumento poderoso en contra de un medio tan extendido de tipo "radiante": es casi seguro que habría impedido la formación de galaxias y de cúmulos de galaxias.

En principio puede determinarse la densidad de toda la materia del universo, esté o no asociada con las galaxias, por comparación del estado presente del universo con el que tenía minutos después de la gran explosión. Las hipótesis



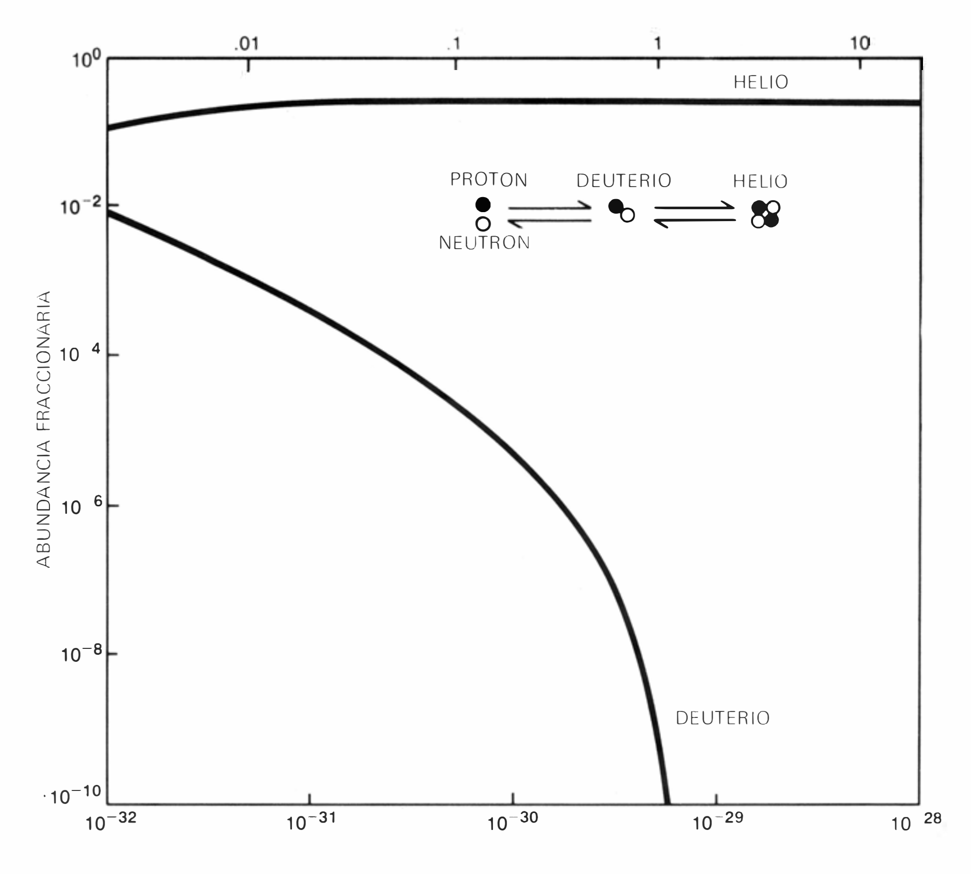
LAS LIGADURAS INICIALES sobre el estado del universo se deducen a partir de la determinación de su edad y del parámetro de deceleración. Los cálculos de la edad de las estrellas más viejas y de la edad media de los elementos pesados sugieren que el universo tiene de ocho a dieciocho mil millones de años; los límites que imponen al tiempo de Hubble dependen de la densidad. Mediante observación de las galaxias lejanas se obtiene un límite superior para el parámetro de deceleración, que no será mayor que 2; así el parámetro de densidad no rebasará el 4. Las ligaduras deducidas de estas medidas no determinan por sí solas si el universo es abierto o cerrado pues incluyen ambas clases de modelos.

más sencillas referidas a esta épocaprimitiva sugieren que la temperatura y la densidad eran suficientemente altas como para que interaccionaran algunas partículas subatómicas y formaran cantidades considerables de algunos núcleos ligeros. En particular, un protón y un neutrón podrían fusionarse, originando un núcleo de deuterio; la mayoría de núcleos de deuterio se combinarían rápidamente para formar núcleos de helio, que consta de dos protones y dos neutrones. La proporción de deuterio y helio formado así depende de la densidad que poseía el universo cuando su temperatura era tal que podía provocar las reacciones. A partir de la densidad primitiva y de la temperatura actual de la radiación de fondo de microondas cabe deducir la densidad presente.

La densidad primitiva

Los modelos matemáticos muestran que, para el margen posible de densidades del universo primitivo, hubo de convertirse en helio entre un 20 y un 30 por ciento de la materia. La cantidad de helio medida en múltiples objetos astronómicos se mueve en este estrecho margen, lo que corrobora rigurosamente la hipótesis fundamental de que el universo pasó por un período de temperaturas y densidades muy elevadas, poco después de la gran explosión. La cantidad actual de deuterio depende estrechamente de la densidad primitiva *(véase la ilustración de la derecha)*. El *Copernicus*, tercer satélite de la serie "Orbiting Astronomical Observatory" (observatorio astronómico en órbita), midió la cantidad relativa de deuterio en el espacio interestelar próximo. Después de tener en cuenta el deuterio consumido en las reacciones nucleares en las estrellas, la cantidad medida da una densidad media actual de unos 4×10^{-31} gramos por centímetro cúbico. Esta medida es un indicador sensible de la densidad; si el universo hubiera tenido una densidad diez veces inferior, la gran explosión habría producido menos de la milésima parte de la cantidad de deuterio calculada. Por ello las imprecisiones derivadas de la medición no son graves respecto a la densidad estimada.

El tiempo de Hubble determina que la densidad obtenida a partir de la abundancia del deuterio suponga que el universo sea abierto o cerrado. Como se ha visto, si el tiempo de Hubble es de diecinueve mil millones de años, la densidad crítica será de 5×10^{-30} gramos por centímetro cúbico; es decir, Ω, que representa el cociente entre la densidad actual y la densidad crítica, valdrá más o menos



LA DENSIDAD DEL UNIVERSO PRIMITIVO influyó en la síntesis del deuterio y el helio; la densidad actual puede deducirse de la abundancia relativa de tales elementos. Se cree que el deuterio se formó por fusión de protones y neutrones a los escasos primeros minutos de la gran explosión; si la densidad hubiera sido entonces excesivamente grande, casi todo el deuterio se habría convertido en helio. La cantidad de ambos elementos viene representada en fracción (respecto a la masa) de toda la materia del universo. Si los modelos más sencillos del universo primitivo son correctos y si no se ha formado deuterio en sucesos más recientes, la cantidad observada sugiere que la densidad del universo no puede superar los 4×10^{-31} gramos por centímetro cúbico, aproximadamente.

0,08. Para cualquier valor del tiempo de Hubble entre trece y diecinueve mil millones de años, el valor Ω derivado de la abundancia de deuterio es compatible con el derivado de la densidad de las galaxias. Recíprocamente, para todo el valor razonable del tiempo de Hubble, Ω no puede alcanzar la unidad, pues de lo contrario sería incompatible con la densidad necesaria para la producción de deuterio.

Pudiera parecer que la cantidad de deuterio aporta una prueba evidente del carácter abierto del universo. Por desgracia, los argumentos que apoyan una tal conclusión son algo inseguros. En la extrapolación del estado presente del universo a las condiciones inmediatamente siguientes a la gran explosión, hemos empleado el modelo más sencillo posible; otros modelos podrían implicar que las cantidades de helio y deuterio observadas provienen de un universo mucho más denso y cerrado. Estos modelos son más complicados, no exentos de imaginación, pero tampoco hay que desecharlos. Además, la significación de la cantidad de deuterio depende absolutamente de la hipótesis que sostiene que todo el deuterio del universo aparecería poco después de la gran explosión. Se han insinuado otras fuentes, tales como las supernovas, pero hasta ahora no se ha encontrado ningún mecanismo que creara una cantidad considerable de deuterio sin violar otras ligaduras.

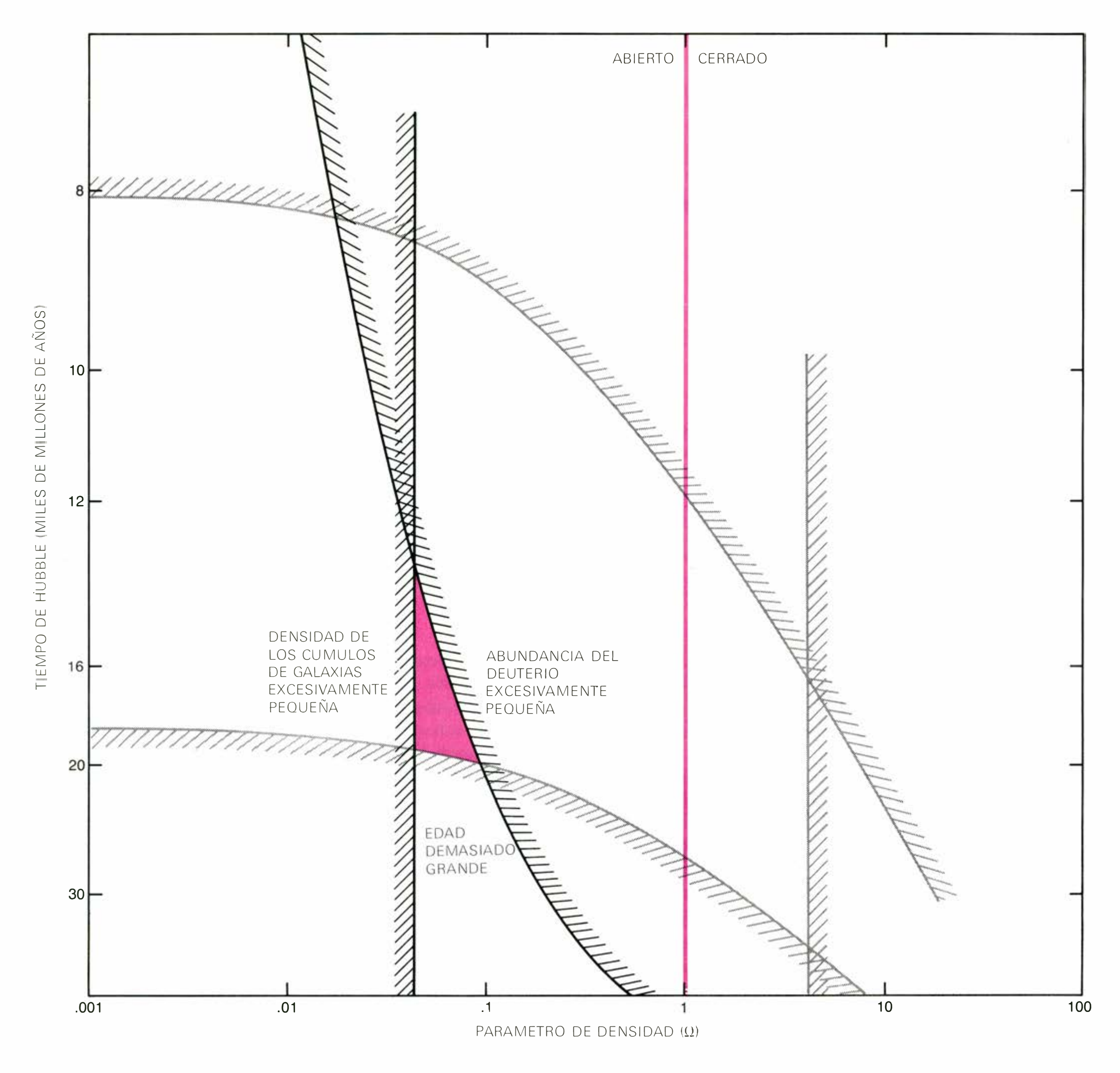
Modelos aceptables

Las mediciones del parámetro de deceleración, de la edad del universo, de la densidad de las galaxias y de la abundancia de deuterio imponen al estado del universo sus propias ligaduras. Si las mediciones fueran compatibles, debería haber algún tipo de modelo de universo que cubriera todas las ligaduras. Tal modelo

existe, y, además, es relativamente pequeño, de forma que pueden aventurarse predicciones válidas sobre el futuro del universo *(véase la ilustración inferior).* Si la edad del universo no resulta excesiva y su densidad es, como mínimo, igual a la observada en las galaxias y no demasiado grande para que se produzca deuterio, el valor de Ω deberá hallarse entonces entre 0,04 y 0,09. Este valor es inferior en mucho al necesario para que el universo sea cerrado.

Hay dos observaciones suplementarias compatibles con los valores permitidos de Ω y del tiempo de Hubble. El cálculo de la edad de las estrellas en los cúmulos globulares es muy sensible a la abundancia de helio, y, como hemos visto, ésta viene determinada a su vez por la densidad del universo. Estimula, pues, observar que la edad y la abundancia de helio permitidas por el conjunto de ligaduras son coherentes con lo que se conoce de las estrellas de los cúmulos globulares.

Las ligaduras reducen el propio tiempo de Hubble a un intervalo entre trece y veinte mil millones de años. La determinación directa del tiempo de Hubble es difícil, pero en los últimos años Allan R. Sangade y Gustav A. Tammann, de los Observatorios de Hale, han venido realizando grandes esfuerzos para lograrlo. Su valor más aproximado es de dieciocho ± dos mil millones de años. Robert P. Kirshner y John Kwan, del Instituto de Tecnología de California, midieron el tiempo de Hubble recurriendo a una técnica distinta: en función



LAS NUEVAS LIGADURAS se combinan mostrando que el universo se expansionara siempre. La cantidad de deuterio pone una cota superior a la densidad de toda la materia del universo, y de ahí que limite también al parámetro de densidad, aunque el valor numérico de esta cota dependa del tiempo de Hubble. El cálculo de la edad de las estrellas y de los elementos pesados da un valor máximo del tiempo de Hubble. Finalmente, la cota inferior de parámetro de densidad queda determinada por la deducción de la masa asociada a un cúmulo de galaxias. Salvo alguna excepción estas ligaduras obligan a todos los modelos a moverse dentro de un pequeño margen de valores respecto al parámetro de densidad y al tiempo de Hubble *(zona coloreada);* **los modelos de esta zona describen un universo abierto, infinito y en expansión perpetua.**

de las propiedades de las estrellas que estallan en galaxias lejanas. Para estos investigadores el valor se sitúa, entonces, entre trece y veintidós mil millones de años.

La coherencia de los resultados obtenidos siguiendo métodos tan diversos es satisfactoria y apoya la confianza en que el modelo cosmológico sea el adecuado y pueda conocerse el futuro del universo. Con todo, el acuerdo que se desprende podría resultar fortuito, habida cuenta de los márgenes de error que se contabilizan en los datos y en la teoría empleada para interpretarlos debidamente.

Un presupuesto común a todos los modelos aquí considerados es que el parámetro de deceleración ha de ser igual a la mitad del parámetro de densidad, hipótesis de trabajo que, según hemos ido viendo, no ha podido comprobarse todavía. Ahora bien, si en el futuro se demostrara que ello es falso se habrían de idear modelos cosmológicos algo más complicados. Tal, una clase de modelos que se basa en una modificación, sugerida por Einstein hace tiempo, de la relatividad general, al introducir una constante llamada constante cosmológica.

En estos modelos aquí considerados, es el propio espacio el que genera una fuerza gravitatoria, que podrá ser atractiva o repulsiva; en consecuencia, la deceleración de la expansión no está relacionada de un modo tan inmediato con la densidad.

Tomadas una a una, todas las ligaduras que hemos analizado tienen posibles escapatorias. En particular, algunos colegas no estarán de acuerdo con la pequeña densidad deducida de la estimación de la masa de las galaxias ni con la limitación a la densidad sacada a partir de la producción de deuterio. Con todo, nuestros argumentos y conclusiones basan su credibilidad en que puede construirse un modelo cosmológico coherente con la interpretación obvia de cada una de las pruebas de evidencia. Debe destacarse que factores tan distintos como la edad de las estrellas, la masa de las galaxias, la abundancia de elementos químicos, y la intensidad observada en la expansión del universo, puedan interpretarse de un modo tan natural según las normas de uno de los modelos cosmológicos más sencillos. Este modelo cosmológico mencionado describe un universo de dimensiones infinitas y en expansión indefinida. La causa a favor de un universo abierto no es determinante, pero se apoya firmemente en el peso de la evidencia.

La búsqueda de nuevas familias de partículas elementales

En experimentos con neutrinos de alta energía ha aparecido una partícula, por lo menos, que no pertenece a ninguna familia conocida. La nueva partícula debe presentar alguna propiedad por descubrir en la materia

David B. Cline, Alfred K. Mann y Carlo Rubbia

Dentro de esa continua interrelación entre descubrimiento experimental e interpretación teórica de la física moderna, el momento actual resulta de marcado interés. Han entrado en funcionamiento nuevas máquinas para crear partículas elementales, y se han inventado nuevos aparatos y nuevas técnicas para detectarlas. Contemporáneamente, las nuevas teorías permiten sospechar que estados de la materia hasta ahora desconocidos podrían descubrirse con estas máquinas. Tal convergencia de teoría y experimentación ha inspirado una búsqueda febril de nuevas familias de partículas elementales, fruto de la cual fue el descubrimiento, en 1974, por investigadores del Brookhaven National Laboratory y del Centro del acelerador lineal de Stanford, de las partículas llamadas *J* o psi. Otra partícula, por lo menos, con propiedades no registradas se ha detectado en experimentos llevados a cabo en el "Fermi National Accelerator Laboratory" (Fermilab). Todavía no se ha podido establecer si las dos nuevas clases de partículas se hallan estrechamente relacionadas.

La partícula observada en el Fermilab tiene una masa y una vida media que la coloca en una región del espectro de partículas elementales prácticamente inexplorada. La partícula se produjo en el Fermilab mediante las interacciones de un haz intenso de neutrinos de alta energía, y su existencia se dedujo indirectamente ante la presencia de neutrinos entre los productos de su desintegración. El experimento dependía críticamente de las propiedades características de los neutrinos; no pudo, pues, realizarse hasta que no existieron los aparatos y las técnicas adecuadas para manejar estas partículas tan elusivas.

La nueva partícula no es una más a añadir a las muchas familias ya conocidas. Su importancia estriba en que demuestra la aparición de una familia totalmente nueva. La existencia de la nueva familia implica, a su vez, que hay alguna propiedad de la materia no observada hasta ahora, que distingue las partículas pertenecientes a esa nueva familia de las otras conocidas. La naturaleza de la nueva propiedad nos sigue siendo oscura, y hay más confusión que certeza en la interpretación de los experimentos del Fermilab. A pesar de lo cual, la física de partículas atraviesa por uno de los momentos más apasionantes de su historia reciente.

Los físicos distinguen cuatro interacciones fundamentales en la materia, o, expresado de otra forma, cuatro fuerzas fundamentales. Reciben el nombre de interacciones fuertes, electromagnéticas, débiles y gravitatorias. Si a las fuerzas fuertes asignamos el valor 1, la intensidad de la fuerza electromagnética será 1/137, aproximadamente 10^{-2}, 10^{-13} la de la fuerza débil y la intensidad de la gravitatoria será 10^{-39}. Cuando dos partículas interaccionan a través de una de estas fuerzas, se cree que la fuerza se transmite por el intercambio de una partícula intermedia que es característica de esta fuerza. La intensidad de la interacción está relacionada con el tiempo necesario para este intercambio y con la distancia a que el intercambio puede realizarse. Tiempo y distancia que, a su vez, se hallan relacionados con las propiedades de la partícula intercambiada.

Las interacciones fuertes suelen ocurrir en unos 10^{-23} segundos; y porque puede originar cambios en intervalos tan breves de tiempo es por lo que a su fuerza se la conoce con la expresión de fuerza fuerte. La interacción electromagnética necesita un tiempo 137 veces mayor aproximadamente, es decir, unos 10^{-21} segundos. El tiempo requerido para los procesos débiles varía considerablemente con la energía de la interacción; desde una levísima duración de 10^{-18} segundos hasta otra prolongada de 15 minutos.

Se cree que el radio de acción de las interacciones fundamentales es inversamente proporcional a las masas de las partículas intercambiadas. Esta relación se ha confirmado para la interacción electromagnética: la fuerza electromagnética es transmitida por fotones, que tienen masa nula, y el radio de acción de la fuerza es infinito. La partícula asociada con la propagación de la fuerza gravitatoria, el gravitón, debe tener masa nula, para poder explicar el radio de acción infinito de la gravedad, pero los gravitones no se han detectado todavía. La fuerza fuerte se transmite por las partículas llamadas mesones, que tienen masa, debido a lo cual la fuerza fuerte posee un radio de acción finito. El mesón menos pesado, el mesón pi o pión, tiene una masa de unos 140 MeV o ciento cuarenta millones de electrón-volts (en la física moderna de altas energías se acostumbra medir la masa dando su equivalente en energía). El alcance o radio de acción de la interacción fuerte es de unos 10^{-13} centímetros, que corresponden aproximadamente al tamaño de las partículas que constituyen el núcleo atómico.

La fuerza débil tiene un alcance muy corto; cuando Enrico Fermi la descubrió por primera vez en 1934, se creyó que su radio de acción era nulo, lo cual habría exigido una partícula intermedia de masa infinita. Teorías más recientes sugieren que la fuerza débil se halla ex-

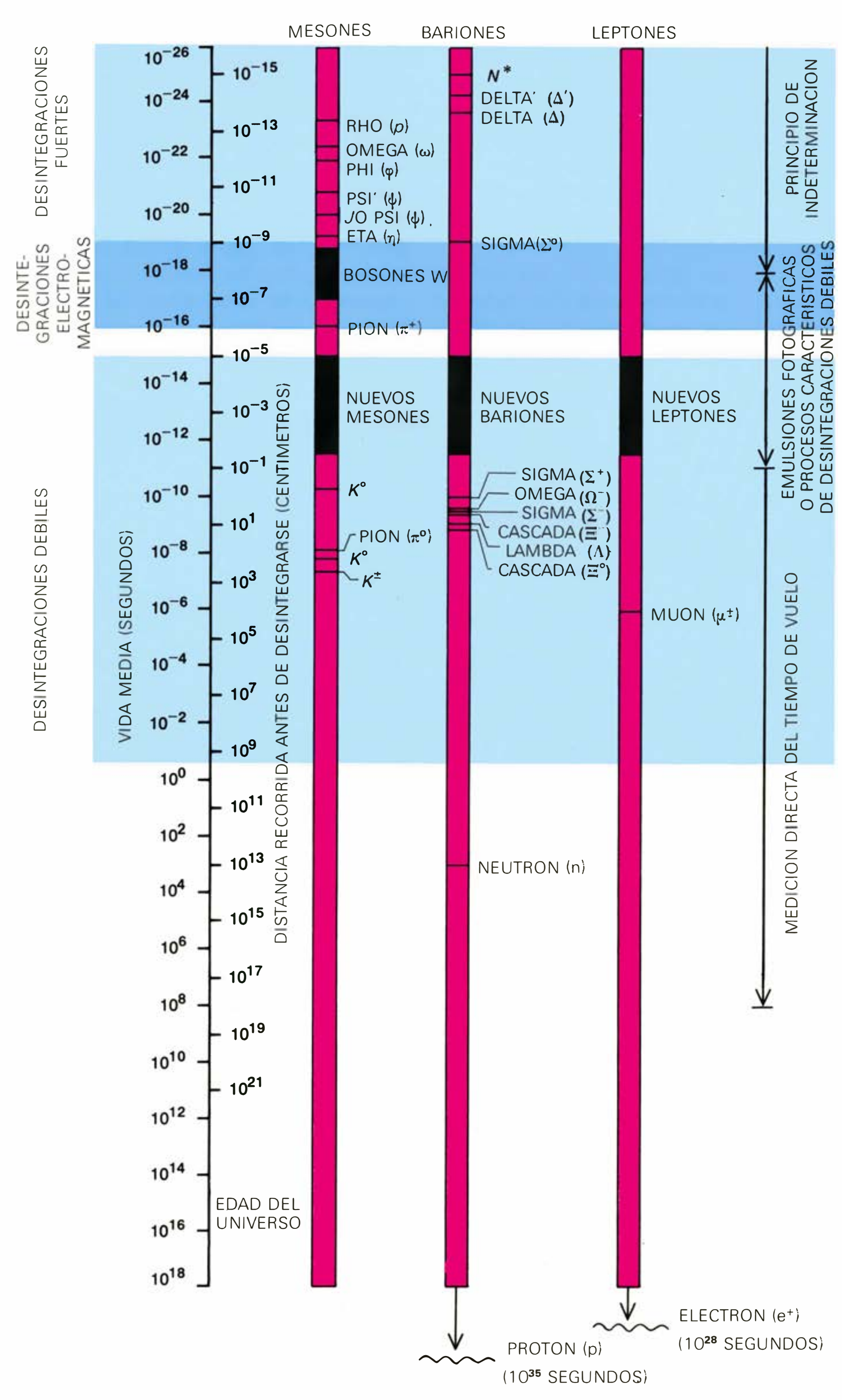
tendida en el espacio, si bien su alcance es extraordinariamente pequeño: menos de 10^{-15} centímetros, una centésima parte del radio de acción de las interacciones fuertes. Se prevé, por tanto, que la partícula intercambiada en las interacciones débiles posea una masa muy grande; mediciones realizadas en el Fermilab sugieren una masa mínima de 20 GeV, o sea, veinte mil millones de electrón-volts. Estas partículas no se han detectado todavía, pero se les dio un nombre; durante muchos años se las ha llamado partículas *W*.

De las cuatro interacciones fundamentales únicamente la gravitatoria es universal. La fuerza débil actúa sobre todas las partículas menos sobre el fotón. La fuerza electromagnética afecta principalmente a las partículas que tienen carga eléctrica. La fuerza fuerte es la más selectiva; sirve de base para dividir las partículas elementales en dos amplios grupos, los hadrones, que "sienten" las interacciones fuertes, y los leptones, que no las "sienten".

Las partículas se clasifican dentro de estos grupos de acuerdo con un sistema muy elaborado de números cuánticos, asignados sobre la base de mediciones experimentales. Los números cuánticos describen las propiedades de una partícula de la misma forma que una serie de rasgos faciales puede describir a una persona. La lista completa de los números cuánticos define la partícula de forma única; junto con la masa, tal lista representa todo lo que posiblemente se puede conocer acerca de la partícula.

Cada número cuántico corresponde a una propiedad que se conserva durante las interacciones de las partículas, es decir, la suma de los números cuánticos relativos a una propiedad determinada no cambia durante la interacción. Algunos números cuánticos, tales como el momento angular de spin o la carga eléctrica, se conservan invariablemente, y lo mismo sucede con la energía y la masa. En la interacción fuerte todos los números cuánticos se conservan absolutamente; en la interacción electromagnética, se conservan todos menos uno (el llamado spin isotópico). En la interacción débil, sin embargo, cambian a veces algunos números cuánticos. Los números cuánticos cuya conservación se altera pertenecen todos a los hadrones y comprenden propiedades tales como la extrañeza, la paridad intrínseca y la conjugación de carga.

El número cuántico más importante para la clasificación de las partículas es el momento angular de spin. Cuando se



LA VIDA MEDIA DE UNA PARTICULA se determina básicamente por su forma de desintegración; la vida media determina, a su vez, la forma en que puede detectarse la partícula. Muchas partículas que pueden desintegrarse sólo mediante la interacción débil tienen vidas medias relativamente largas, y recorren una distancia considerable antes de desintegrarse. Su promedio de vida puede determinarse a través de una medición directa de su camino en un detector de partículas. Las partículas que se desintegran mediante una interacción fuerte o electromagnética sobreviven muy poco tiempo, y sus recorridos, excesivamente cortos, no pueden medirse. Por el contrario, sí puede determinarse su promedio de vida a través del principio de indeterminación de Werner Heisenberg, que dice que la vida media es inversamente proporcional a la indeterminación de la energía, o de la masa, de la partícula. Ninguno de estos métodos es práctico para partículas con promedio de vida intermedia, entre los 10^{-18} segundos y 10^{-11} segundos aproximadamente. Algunas trazas cortas de partículas pueden detectarse en emulsiones fotográficas, pero la mayoría de las partículas, con vidas medias comprendidas en esta región deben detectarse indirectamente, a partir de los productos de su desintegración. Se prevé que varios grupos de partículas pueden estar en esta región *(barras grises)*. Se espera que se desintegren a través de las interacciones débiles, aunque con mayor rapidez que otras partículas que decaen también débilmente.

mide en unidades naturales mediante la constante de Planck dividida por 2π, su valor es siempre un entero o un semi-impar. Las partículas con valores enteros del momento angular de spin reciben el nombre de bosones, porque siguen la llamada estadística de Bose-Einstein. Las partículas con spin semi-impar siguen la estadística de Fermi-Dirac, y reciben el nombre de fermiones.

Todos los leptones conocidos son fermiones; en efecto, todos ellos tienen el mismo momento angular de spin igual a 1/2. Por otra parte, entre los hadrones hay bosones y fermiones. Los bosones hadrónicos son los mesones (las partículas que transmiten la interacción fuerte), y tienen valores del spin iguales a 0, 1, 2 o 3. Los fermiones forman un grupo distinto, llamado de los bariones, con números cuánticos de spin iguales a 1/2, 3/2, 5/2, etc. Dentro de cada uno de los grandes grupos de hadrones las partículas se reúnen en familias, que se distinguen por su valor común del momento angular de spin.

El fotón y las partículas *W* no son ni hadrones ni leptones. Tienen spin 1, por tanto serán bosones; se las suele designar con el nombre de bosones vectoriales porque la ecuación mecanicocuántica que describe las partículas de spin 1 toma la forma de un vector tetradimensional. Los gravitones, con spin 2, son también bosones, y lo mismo sucede con los mesones que transmiten la interacción fuerte. Así pues, todas las fuerzas fundamentales se propagan por bosones.

			INTERACCION FUERTE	INTERACCION ELECTROMAGNETICA	INTERACCION DEBIL
		PARTICULA INTERMEDIA	MESONES	FOTON (γ)	W^+ W^- W^o
		MASA DE LA PARTICULA INTERMEDIA	140 MeV Y MAYOR	0	MAYOR QUE 20 GeV
		ALCANCE	10^{-13} CENTIMETROS	INFINITO	10^{-15} CENTIMETROS
		ESCALA DE TIEMPO	10^{-23} SEGUNDOS	10^{-21} SEGUNDOS	10^{-10} SEGUNDOS
		ENERGIA	CONSERVADO	CONSERVADO	CONSERVADO
		MOMENTO	CONSERVADO	CONSERVADO	CONSERVADO
		CARGA ELECTRICA	CONSERVADC	CONSERVADO	CONSERVADO
		NUMERO LEPTONICO (L_e, L_μ)	CONSERVADO	CONSERVADO	CONSERVADO
		EXTRAÑEZA	CONSERVADO	CONSERVADO	NO CONSERVADO
		ENCANTO	CONSERVADO	CONSERVADO	NO CONSERVADO
HADRONES	BARIONES	PROTON (p)			
		NEUTRON (n)			
		LAMBDA (Λ)			
	MESONES	PION ($\pi^\pm$)			
		PION (π^o)			
		KAON ($K^\pm$)			
		KAON (K^o)			
LEPTONES		ELECTRON ($e^\pm$)			
		MUON ($\mu^\pm$)			
		NEUTRINO (ν_e, ν_μ)			

LAS FUERZAS FUNDAMENTALES, o interacciones entre partículas, se transmiten por partículas intermedias, según se cree. La escala de los tiempos de la interacción es una medida de su intensidad, en tanto que el alcance de la interacción es inversamente proporcional a la masa de la partícula intermedia. Algunas propiedades, tales como la energía, el momento y la carga eléctrica, se conservan en todas las interacciones; otras, definidas por números cuánticos tales como la extrañeza, no se conservan invariablemente en las interacciones débiles. El encanto es un número cuántico hipotético, análogo a la extrañeza. La interacción fuerte afecta sólo a la clase de partículas llamadas hadrones, esto es, a los bariones y mesones. El electromagnetismo afecta a todas las partículas con carga eléctrica. La fuerza débil viene "sentida" por todas las partículas, menos por el fotón.

Los bariones con masa menor tienen todos como valor de spin 1/2. Hay ocho: dos nucleones *(N)* con una masa media de 939 MeV, una partícula lambda (Λ) con una masa de 1.116 MeV, tres partículas sigma (Σ) con una masa media de 1.193 MeV y dos partículas csi (Ξ) con masa media de 1.318 MeV. ¿Cómo se explican estas diferencias de masa? Suelen atribuirse a una hipotética quinta interacción, llamada corrientemente fuerza semifuerte. Esta quinta fuerza no tiene ni la importancia teórica ni la evidencia experimental de las cuatro fuerzas fundamentales, y se introduce únicamente para explicar los "desdoblamientos másicos" de los hadrones. Si se pudiera eliminar de alguna forma la interacción semifuerte, todos los bariones de spin 1/2 que acabamos de mencionar tendrían la misma masa y formarían un solo estado degenerado. La fuerza semifuerte desdobla el estado degenerado en partículas de distintas masas, desdoblamiento que viene acompañado por la introducción de dos nuevos números cuánticos, spin isotópico y extrañeza. Ambos son números cuánticos "internos": no pueden observarse directamente sus efectos en el comportamiento de una partícula individual, sino que son consecuencia de las características de las familias de hadrones.

A través de la interacción semifuerte, partículas con valores distintos de la extrañeza o del spin isotópico adquieren masas distintas *(véase la ilustración superior de la página opuesta).* Las partículas lambda y sigma, por ejemplo, tienen ambas extrañeza –1, pero como tienen valores diferentes del spin isotópico (0 y 1, respectivamente) presentan una diferencia de masa de unos 77 MeV. La cascada (csi) y el nucleón tienen ambos el mismo spin isotópico pero su extrañeza difiere en dos unidades; como consecuencia, la masa de la partícula csi es superior a la del nucleón en unos 375 MeV.

Con un mecanismo similar la interacción electromagnética desdobla cada estado hadrónico en partículas de masa y carga eléctrica distinta (y por lo tanto

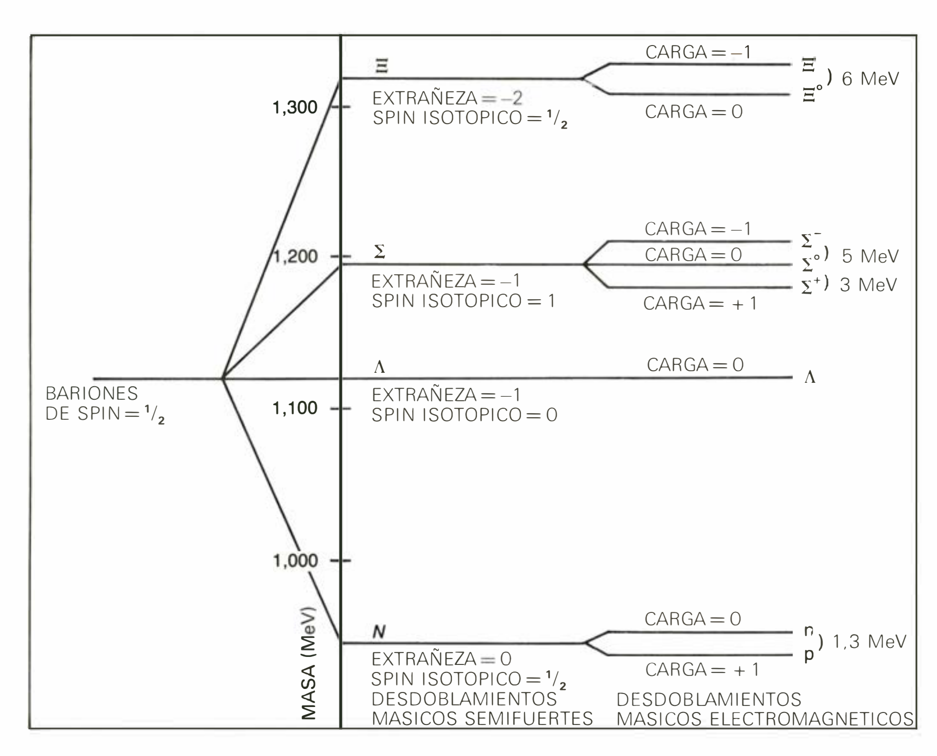
viola la conservación del spin isotópico). Si se pudiera eliminar la fuerza electromagnética, los dos nucleones, por ejemplo, tendrían idéntica masa, y lo mismo sucedería con otras propiedades suyas. Como no hay forma de eliminar la fuerza, el nucleón se desdobla en dos estados, el protón con carga eléctrica + 1 y el neutrón con carga eléctrica 0. Al ser la fuerza electromagnética menos intensa que la semifuerte, la diferencia de masas entre el neutrón y el protón es sólo de 1,3 MeV, unas 136 veces menor que la diferencia de masa entre la partícula lambda y el nucleón.

Los mismos mecanismos generan desdoblamientos másicos de igual orden de magnitud entre los mesones *(véase ilustración inferior de esta misma página).* Los mesones de menor masa son los que tienen momento angular de spin igual a 0; éstos comprenden tres piones (π), cuatro kaones (K), o mesones K y los mesones eta (η) y eta' (η'). Los estados pión y K tienen spin isotópico y extrañeza distintos, siendo la diferencia de masa entre ambos de unos 350 MeV. Las mesones eta y eta' constituyen un caso especial: tienen idénticos valores de spin isotópico y extrañeza, pero difieren en otro número cuántico, y tienen una diferencia de masa de unos 410 MeV. La interacción electromagnética desdobla los piones en estados con carga eléctrica 0, + 1 y –1 y con una diferencia de masa de unos 4 MeV. Los desdoblamientos másicos de origen electromagnético de los kaones son más complicados, por tratarse éstos de partículas extrañas. Además de kaones en estado positivo y negativo, hay dos neutros con valores de extrañeza opuestos.

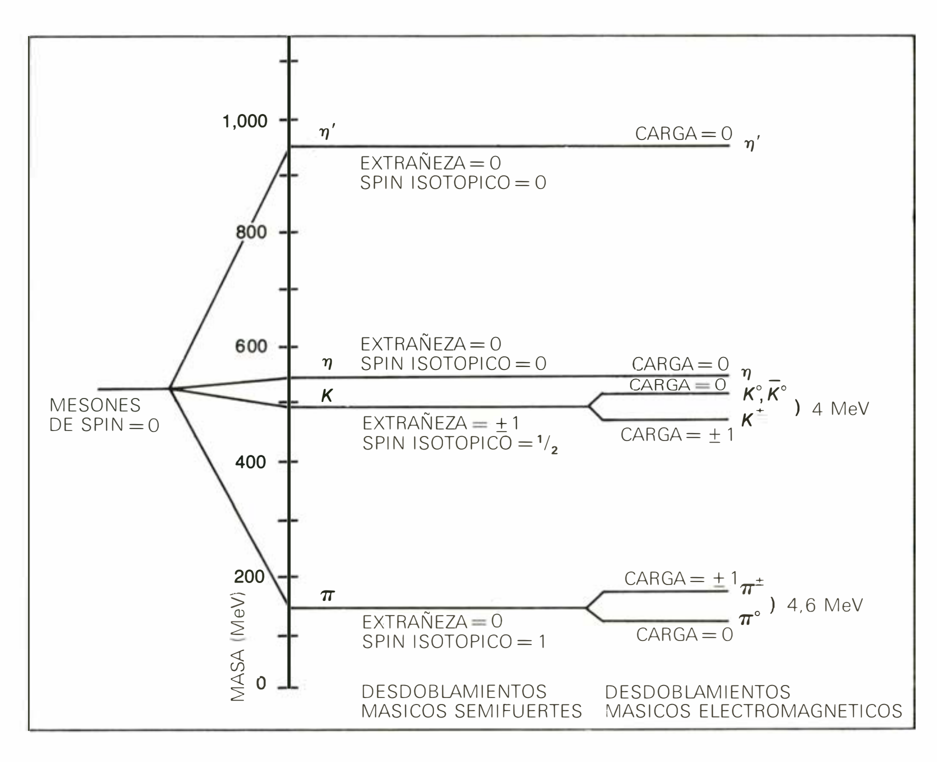
Los desdoblamientos másicos aparecen también entre los leptones, pero no existe una explicación satisfactoria para los mismos. La familia de los leptones contiene únicamente cuatro partículas: el electrón *(e),* el muón *(μ)* y dos tipos de neutrinos (ν_e y ν_μ).

Los electrones son extraordinariamente ligeros, con una masa de unos 0,51 MeV; el muón pesa unas 200 veces más que el electrón, con una masa de unos 105,6 MeV. No se conoce ningún mecanismo que explique este desdoblamiento másico; más aún, la misma existencia del muón es uno de los misterios de más larga historia en la física de partículas.

El electrón se contradistingue del muón por un número cuántico leptónico característico; se designan, respectivamente por L_e y L_μ. Estos números cuánticos se conservan en todas las interacciones y van transportados por las dos clases de neutrinos. Los neutrinos, al carecer de



LAS MASAS DE LOS BARIONES se explican postulando la existencia de un estado único hipotético, que se "desdobla" en partículas de masa distinta. Los bariones con momento angular de spin igual a 1/2 comprenden cuatro clases de partículas: el nucleón *(N),* la lambda (Λ), la sigma (Σ) y la csi (Ξ). Ellos sufren una diferencia de masa debida a una hipotética interacción semifuerte, en razón de los valores de los números cuánticos de extrañeza y de spin isotópico. Cada partícula de éstas se desdobla luego, mediante una interacción electromagnética, en estados de cargas eléctricas distintas. El nucleón por ejemplo, se divide en neutrón y protón. Los desdoblamientos másicos electromagnéticos, que son más pequeños que los semifuertes, no están a escala.



LOS DESDOBLAMIENTOS MASICOS DE LOS MESONES son mayores que los de los bariones, pero están generados por el mismo mecanismo. Los mesones con momento angular de spin igual a cero se desdoblan en cuatro estados mediante la hipotética interacción semifuerte: el pión (π), el kaón (K), la eta (η) y la eta' (η'). El pión y el kaón se desdoblarán ulteriormente, a través de las interacciones electromagnéticas, en partículas neutras y cargadas. Para partículas que se desintegran a través de las interacciones débiles, el tamaño del desdoblamiento másico entre los estados influye en la vida media de las partículas: cuanto mayores sean los desdoblamientos másicos, más rápidas serán las desintegraciones. De acuerdo con el estado de las investigaciones se espera que las nuevas familias de partículas tengan desdoblamientos másicos mayores y, por tanto, desintegraciones más rápidas que los mesones y bariones conocidos.

carga eléctrica, no sienten la fuerza electromagnética; en cuanto leptones, se muestran insensibles a la fuerza fuerte. Los neutrinos sólo pueden interaccionar mediante las fuerzas débiles y, en consecuencia, apenas interaccionan con la materia. Se cree que no tienen masa, si bien este hecho no se ha confirmado por la experiencia.

Los mayores desdoblamientos másicos son los previstos para los bosones vectoriales. Como quedó expuesto, la diferencia de masa entre el fotón (de masa nula) y las partículas W debe ser de 20 GeV por lo menos, pudiendo ser mucho mayor. Los bosones cargados eléctricamente, W^+ y W^-, fueron introducidos por Hideki Yukawa en 1935; el bosón neutro, W^0, se incorporó a esta familia en la década de los sesenta, y condujo a la visión de corrientes débiles neutras, un tipo de interacción en el que las partículas interaccionan mediante fuerzas débiles sin transferencia de carga eléctrica. Recientemente se han observado corrientes débiles neutras pero no pueden considerarse como prueba concluyente en favor de la existencia del W^0.

Previsiones muy próximas a nosotros atribuyen a los bosones W^+ y W^- masas que pueden alcanzar los 54 GeV y, hasta 80 GeV, la masa del W^0. Tales desdoblamientos másicos, enormes, vienen postulados por teorías muy ambiciosas que pretenden demostrar que la fuerza débil y la fuerza electromagnética, no obstante sus aparentes diferencias, son en puridad dos manifestaciones distintas de un mismo fenómeno. Estas teorías explican los desdoblamientos másicos suponiendo que, en un hipotético estado inicial, los cuatro bosones vectoriales carecen de masa; de lo cual se sigue consecuentemente que tres de estas partículas adquieren una masa enorme, mientras que la cuarta mantiene su masa nula.

Las teorías unificadas de las interacciones débiles y electromagnéticas sugieren también la posible existencia de nuevos leptones, que tendrían masas mucho mayores que los conocidos hasta ahora. Tanto los bosones pesados W como los leptones pesados suplementarios desempeñan la misma función en las teorías, a saber: eliminar ciertos términos matemáticos infinitos que, en caso contrario, destruirían el significado y el valor predictivo de las teorías.

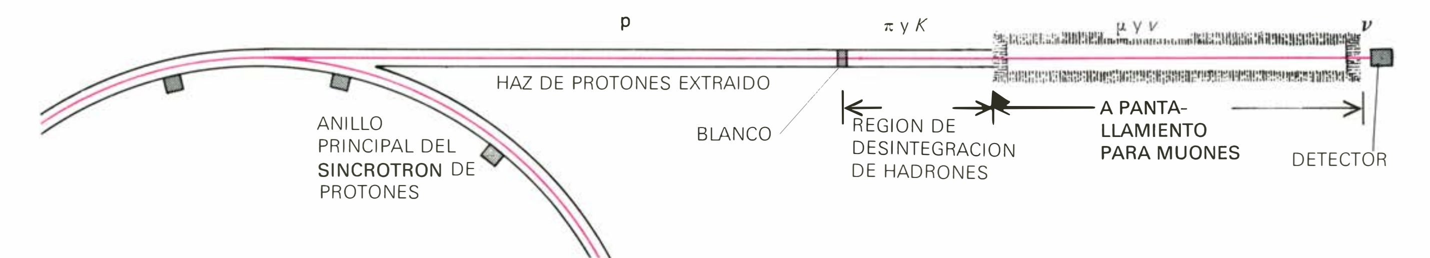
De las muchas partículas identificadas, sólo unas pocas son estables, y constituyen las componentes de la materia ordinaria. Se sabe que el fotón es estable; los principios fundamentales de la mecánica cuántica sostienen que no puede desintegrarse. Se cree que el protón y el electrón Son estables, suposición que se basa en los datos experimentales que indica que sus vidas medias deben ser muchos órdenes de magnitud mayores que lo que se estima sea la edad del universo. Por último, se cree que los neutrinos son también estables, pero no se ha demostrado rotundamente.

Las demás partículas, por tender al estado de energía más baja, decaen más o menos rápidamente, alcanzando por último un estado final consistente de partículas estables. Algunos hadrones decaen a través de las interacciones fuertes. El mesón eta', por ejemplo, puede decaer en un mesón eta y dos piones de carga opuesta, reacción cuya formalización simbólica es: $\eta' \rightarrow \eta + \pi^+ + \pi^-$. Se comprueba fácilmente que la masa se conserva en esta interacción: el eta' tiene una masa de unos 958 MeV, y la suma de las masas de los productos de la desintegración es de unos 829 MeV, con una pérdida de 129 MeV aprovechados en forma de energía cinética de los productos. La carga eléctrica es conservada también: el eta' es neutro, las cargas de los productos de desintegración son 0, +1 y −1, resultando una carga total nula. De igual forma puede probarse que el momento angular de spin, la extrañeza y el spin isotópico se mantienen. La desintegración tiene lugar en el intervalo de tiempo característico de la interacción fuerte, unos 10^{-23} segundos.

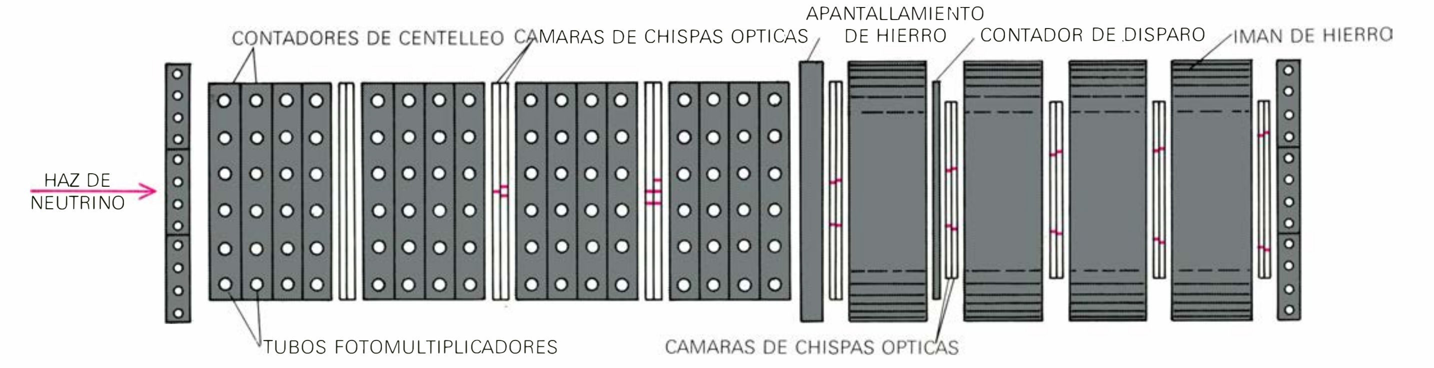
Muchas partículas no pueden desintegrarse a través de las interacciones fuertes al no existir combinaciones de partículas que tengan masa menor y conserven todos los números cuánticos. Algunos hadrones, como el pión neutro, se desintegran electromagnéticamente. El pión neutro es el hadrón de masa más pequeña, y no puede desintegrarse a través de la fuerza fuerte porque no hay partículas con interacción fuerte de masa menor en las que pueda transformarse. De ahí que se desintegre, casi siempre, en dos fotones. Este proceso conserva también la masa, la carga eléctrica y todos los otros números cuánticos. Requiere un tiempo de unos 10^{-16} segundos.

Muchas partículas no pueden seguir ninguno de estos dos procesos, y han de desintegrarse a través de las interacciones débiles. Consideremos, por ejemplo, la desintegración de la partícula lambda, que pertenece a la familia de los bariones, según se dijo, con momento angular de spin 1/2. La partícula lambda se desintegra, con mayor frecuencia, en un protón y un pión negativo o en un neutrón y un pión neutro; estas reacciones se escriben así: $\Lambda \rightarrow p + \pi^-$ y $\Lambda \rightarrow n + \pi^0$. Todas las partículas que intervienen son hadrones, por tanto podría pensarse que la desintegración tiene lugar mediante la interacción fuerte. Se conserva la masa, la carga eléctrica y el momento angular de spin. Pero debe desecharse una desintegración fuerte, porque las interacciones fuertes conservan la extrañeza. El lambda tiene extrañeza −1; el protón, el neutrón y el pión, nula. En ninguna de esas formas de desintegración se conserva la extrañeza sino que cambia en una unidad. La desintegración debe ocurrir, por consiguiente, mediante la interacción débil, siendo el tiempo necesario mucho mayor que para las otras desintegraciones, esto es: unos 10^{-10} segundos.

Hay otros muchos hadrones que deben desintegrarse mediante la fuerza débil, ya que en su desintegración queda violada alguna ley de conservación. Por



HACES DE NEUTRINOS: vienen generados indirectamente a partir de protones acelerados en el sincrotón de protones de 400 GeV del "Fermi National Accelerator Laboratory" (Fermilab). Al colisionar los protones con un blanco metálico se generan los piones y los kaones. A continuación, estos mesones se irán desintegrando principalmente en muones, neutrinos y antineutrinos. Los muones y los hadrones que aún pudieran quedar se eliminan provocando que el haz atraviese un montículo de tierra, de un kilómetro de longitud, aproximadamente, situado al final de la región de desintegración. Cuando el haz sale del montículo, en su composición sólo restan neutrinos y antineutrinos.



EL DETECTOR empleado en el Fermilab para la búsqueda de nuevas familias de partículas es sensible a sucesos con dos muones en el estado final (sucesos dimuónicos). Neutrinos energéticos interaccionan con nucleones en tanques que contienen unas 60 toneladas de un líquido que centellea, es decir, que emite destellos luminosos, cuando una partícula cargada eléctricamente la cruza; los centelleos se registran mediante tubos fotomultiplicadores. Detrás de los contadores de centelleos hay una pantalla de acero que elimina los hadrones y constituye un blanco adicional para la interacción de neutrinos. Las cámaras de chispas ópticas registran los recorridos de las partículas; los imanes las desvían, de forma que pueda medirse su momento. Se cree que las partículas observadas en las últimas cámaras de chispas son muones, porque éstos son las únicas partículas cargadas capaces de atravesar la pantalla de acero y los imanes de hierro. La presencia de dos muones indica que también ha sido creado un neutrino, que no puede detectarse; uno de los muones, por lo menos, y el neutrino podrían producirse por alguna nueva partícula que se desintegre a través de la interacción débil. Esta interacción tiene un alcance muy pequeño.

ejemplo, los kaones dotados de carga que tienen extrañeza ± 1, pueden desintegrarse en un pión cargado y en otro neutro, o en un muón y un neutrino. Al no conservarse la extrañeza, la desintegración habrá de producirse por una interacción débil, y necesitará un tiempo de unos 10^{-8} segundos.

En algunos casos, la desintegración débil es obligada aunque se conserven todos los números cuánticos. Sea la desintegración del neutrón; no puede dar origen a un protón y un pión porque la suma de las masas de estas dos partículas es mayor que la del neutrón; tampoco puede decaer electromagnéticamente en un protón y un fotón porque la reacción violaría la conservación de la carga eléctrica. La única transición permitida es la mediada por las interacciones débiles, y da un protón, un electrón y un antineutrino. Se escribe: $n \rightarrow p + e^- + \bar{\nu}_e$. Este proceso se conoce con el nombre de desintegración beta, es una forma familiar de radiactividad y constituye una parte esencial de la química nuclear de las estrellas.

En la desintegración del neutrón no se viola ninguna ley de conservación. La energía, la carga eléctrica y el momento angular de spin, así como también la extrañeza, son las mismas antes y después de la desintegración. Sin embargo, el tiempo necesario para la desintegración es extraordinariamente largo: 10^3 segundos, unos 16 minutos aproximadamente. Una desintegración muy parecida del barión lambda, $\Lambda \rightarrow p + e^- + \bar{\nu}_e$, es mucho más rápida, aunque viola la ley de conservación de la extrañeza. ¿Por qué está inhibida la desintegración del neutrón?

La explicación comporta un principio importante, aplicable a otras muchas desintegraciones débiles. El desdoblamiento másico entre los estados lambda y nucleón (generado por una interacción semifuerte) es de unos 177 MeV; lo cual representa la energía disponible para que ocurra ese proceso de desintegración. Por otra parte, el desdoblamiento másico entre el neutrón y el protón (generado por una interacción electromagnética) es sólo de 1,3 MeV aproximadamente. Esta falta de energía frena la desintegración del neutrón; y lo hace en mayor medida que la no conservación de la extrañeza puede frenar la desintegración de la partícula lambda.

Este principio demuestra que el desdoblamiento másico en el seno de una familia de partículas se halla estrechamente relacionado con la rapidez de las desintegraciones débiles de las partículas. Si los estados de partícula tienen masas muy próximas, entonces hay muy poca energía disponible para sus desintegraciones y podemos esperar que tengan promedios de vida relativamente largos.

Supongamos que haya un nuevo grupo de hadrones, cuyos desdoblamientos másicos sean mayores que los de las familias conocidas; verbigracia: los bariones de spin 1/2. Las partículas de este grupo habrían de poseer necesariamente un número cuántico nuevo. En efecto, el nuevo número cuántico es lo que definiría al nuevo grupo y distinguiría a sus miembros de las partículas conocidas. Debería tratarse de otro número cuántico interno, tal vez con estrechas analogías con la extrañeza.

Dentro del nuevo grupo, una partícula podría decaer en otra mediante la fuerza fuerte, igual que una partícula extraña puede transformarse en otra (de masa menor) por ese mismo tipo de fuerza. Las transiciones a otros estados, sin embargo, tales como las desintegraciones en bariones y mesones conocidos, tendrían lugar solamente mediante la fuerza débil, porque tales transiciones no conservarían el nuevo número cuántico.

A causa de los grandes desdoblamientos másicos entre nuevos hadrones hipotéticos, tales transiciones débiles no quedarán inhibidas por falta de energía. Por tanto, puede predecirse que una característica típica de estos nuevos hadrones estribaría en unos tiempos de interacción débil considerablemente más cortos que los de las partículas mejor conocidas. El promedio de vida, sin embargo, superaría los 10^{-23} segundos, característicos de las partículas que decaen mediante las interacciones fuertes.

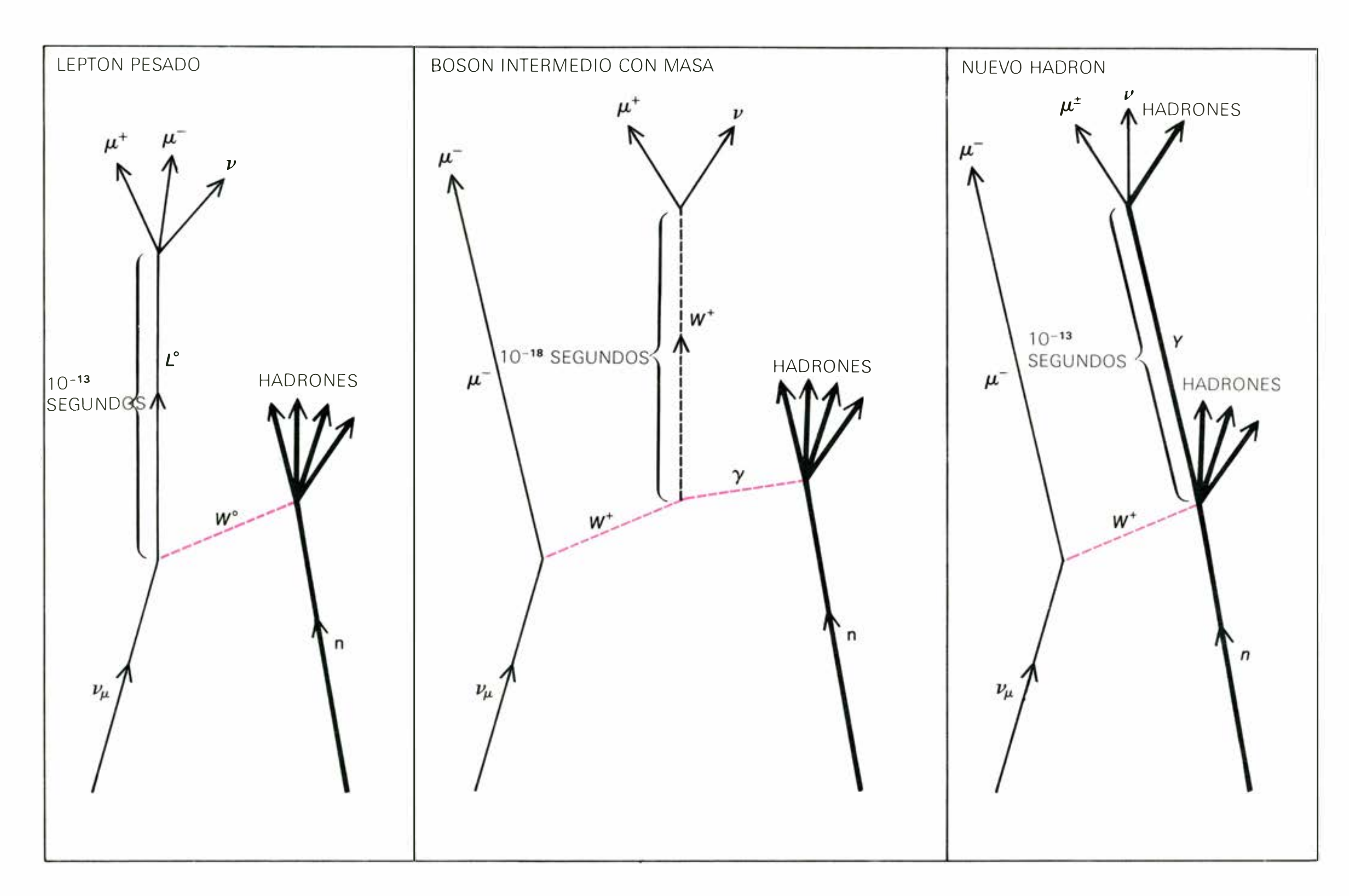
Las desintegraciones de los leptones han de ocurrir necesariamente a través de las interacciones débiles. El electrón y los neutrinos no se desintegran; las desintegraciones del muón, único leptón que es capaz de transiciones a un estado de energía inferior, deben conservar, además de la energía y la carga eléctrica, los números cuánticos leptónicos, L_e y L_μ. En consecuencia, el muón se desintegra siempre en un electrón, un antineutrino de tipo electrónico y un neutrino de tipo muónico. El proceso de desintegración se escribe de la siguiente forma: $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$. El estado inicial tiene un número cuántico L_μ igual a $+1$, y lo mismo sucede con el neutrino muónico del estado final; dado que los números cuánticos L_e del electrón y del antineutrino se eliminan, los números cuánticos leptónicos se conservan. El desdoblamiento másico entre el muón y el electrón es de unos 105 MeV y de unos 10^{-6} segundos el tiempo necesario para la desintegración.

Si existen leptones más pesados, se desintegrarán probablemente de forma análoga al muón, esto es, emitiendo un neutrino y un antineutrino. Pero una vez más, a mayor desdoblamiento másico, más rápida desintegración; es de esperar que los leptones pesados tengan una vida media mucho más corta que el muón.

Los bosones vectoriales intermedios, dotados de masa, deberían desintegrarse también a través de la interacción débil. Una forma probable de desintegración de las partículas *W* cargadas sería por emisión de un electrón o un muón y un antineutrino. Así, por ejemplo, para los bosones vectoriales negativos las reacciones serían $W^- \rightarrow e^- + \nu_e$ y $W^- \rightarrow \mu^- + \nu_\mu$. En ambos casos se conserva la carga eléctrica y los números cuánticos leptónicos L_e y L_μ. Los desdoblamientos másicos que separan las partículas *W* de los leptones son muy grandes; por consiguiente, las desintegraciones débiles deberían ser sumamente rápidas. Se estima que la vida media de las partículas *W* alcanzará los 10^{-18} segundos, aproximadamente.

Hemos ofrecido argumentos en pro de la existencia de tres nuevas clases de partículas, no descubiertas todavía: un nuevo grupo de hadrones, los leptones pesados y los bosones vectoriales intermedios dotados de masa. Todos tienen su interés, desde un punto de vista teórico. Primero, el descubrimiento de un nuevo grupo de hadrones: al introducir un nuevo número cuántico, el nuevo grupo podría ayudar a decidirse sobre la estructura de los hadrones entre las distintas teorías existentes. El descubrimiento y la explicación de las partículas extrañas, hace 25 años, condujo a una nueva síntesis de este tipo. En segundo lugar, los leptones: al constituir una familia reducida, encontrar otro sería un hecho a destacar. Hay más, existen pruebas de que los leptones sean partículas simples, puntuales, sin estructura, y, hasta el presente, nadie aportó una explicación teórica de cómo se originan los desdoblamientos másicos para tales partículas, o por qué se necesitan dos números cuánticos (L_e y L_μ) para describir cuatro partículas. El conocimiento de nuevos estados leptónicos y de sus propiedades podría constituir un primer paso para contestar a estas preguntas. Finalmente, la existencia de las partículas *W*, o de leptones pesados, o de ambos, es esencial para el éxito de las teorías unificadas de las interacciones débiles y electromagnéticas. Si se encuentran las partículas *W*, quedaría confirmada la sugestiva hipótesis de que cada una de las fuerzas fundamentales se transmite por un bosón intermedio, cuya masa es inversamente proporcional al radio de acción de la fuerza.

Supongamos, metodológicamente, que los tres tipos de nuevas partículas existen en la naturaleza. ¿Por qué no se han observado experimentalmente? Pudiera ser que tuvieran masas muy altas y los aceleradores existentes no alcanzaran las energías necesarias para su producción. Si las partículas *W* pesaran lo previsto por ciertas teorías, entonces no se podrían producir con las máquinas de hoy día, pero resulta mucho menos probable que los nuevos hadrones y los leptones pesados posean masas tan elevadas. La alta energía no es, sin embargo, la única limitación de los aparatos de experimen-



LOS SUCESOS DIMUONICOS pueden originarse de tres posibles interacciones entre un neutrino de tipo muónico y un neutrón; en cada una de ellas, el estado final comprende dos muones y un neutrino. La interacción puede crear un leptón pesado neutro (L^o), un bosón con masa intermedio (W^+), uno de los portadores de la fuerza débil, o un hadrón (Y) que tiene un número cuántico nuevo. Las características de los dimuones detectados en el Fermilab sugieren que la interpretación más probable es que se trate de un nuevo hadrón. En estos diagramas, los leptones se representan mediante líneas de trazo finas, los hadrones por líneas de trazo grueso y los mesones intermedios por líneas de puntos. Por su parte, las partículas virtuales, que nunca pueden detectarse, se indican en estos diagramas por unas líneas coloreadas.

tación. No sólo hay que crear una partícula, hay que detectar también su presencia, en cuyo proceso de detección desempeña un papel importante su promedio de vida.

Una partícula que es estable o con una vida media relativamente larga puede identificarse con facilidad. Si está cargada eléctricamente, dejará una traza medible en una cámara de burbujas o en un detector similar. Se podrán medir la velocidad, el momento y la energía. Si es eléctricamente neutra, su detección resultará más difícil pero los productos de sus colisiones pueden identificarse en general.

Estos métodos de detección son útiles para partículas cuyo promedio de vida sea notable, porque exigen que éstas atraviesen una distancia macroscópica en el interior del detector. Las partículas de alta energía suelen viajar a una velocidad próxima a la de la luz; si la partícula tiene una vida media de 10^{-13} segundos, no recorrerá una décima de milímetro antes de desintegrarse. Esta distancia es demasiado pequeña para que se la pueda medir en la mayoría de los detectores, cuya utilidad se limita, por consiguiente, a partículas cuyo promedio de vida supere los 10^{-13} segundos.

Por métodos que son menos directos y que aportan menos información sobre la partícula pueden detectarse también partículas cuyo promedio de vida sea muy corto. El método de detección que se emplea es una aplicación del principio de indeterminación de Werner Heisenberg; consiste en buscar un incremento de la probabilidad de interacción entre partículas conocidas a una cierta energía. Tal incremento recibe el nombre de una resonancia. Hay una expresión del principio de indeterminación que relaciona la indeterminación de la energía a que se produce la resonancia con la vida media de la partícula que se ha producido: a mayor indeterminación, mayor promedio de vida. Muchos hadrones que se desintegran a través de una interacción fuerte en unos 10^{-23} segundos se detectaron por este método. Las partículas psi o J se descubrieron, asimismo, como resonancias. En el Stanford Linear Accelerator Center (Centro del acelerador lineal de Stanford) la partícula menos pesada de todas ellas apareció como un incremento de la probabilidad de que electrones y positrones se aniquilen entre sí a una energía de 3,1 GeV. La incertidumbre en la determinación de tal energía era de varios miles de electrón-volts, lo que permitió cifrar su promedio de vida en unos 10^{-20} segundos. De forma análoga se midió la vida media del mesón eta, que se considera de unos 10^{-19} segundos.

En principio cualquier partícula puede detectarse como una resonancia, pero el método no es práctico para las que tienen un promedio de vida superior a los 10^{-18} segundos. Para partículas de vida media larga, la indeterminación teórica en la energía de la resonancia es muy pequeña, hasta el punto de ser netamente inferior a los errores de medida introducidos por los aparatos experimentales.

Hay un intervalo en el espectro de los promedios de vida no fácilmente medible: se halla comprendido entre los 10^{-18} y los 10^{-11} segundos *(véase la ilustración de la pág. 21)*. Esta región no está completamente vacía; contiene, según hemos mencionado, el pión neutro, con una vida media de 10^{-16} segundos. Sin embargo, salvo esta excepción, la región se encuentra totalmente inexplorada. Si se desea proseguir una investigación sistemática de nuevas partículas, habrá que recurrir a otros métodos de detección.

Se supone que los nuevos hadrones, los leptones pesados y las partículas W tendrán todos un promedio de vida comprendido entre los 10^{-12} y los 10^{-18} segundos. Esta opinión se basa en los desdoblamientos másicos entre los estados de las partículas, que se espera que sean mayores que los de las partículas conocidas. Por consiguiente, no es probable que los métodos convencionales de detección sean eficaces en la búsqueda de los nuevos estados.

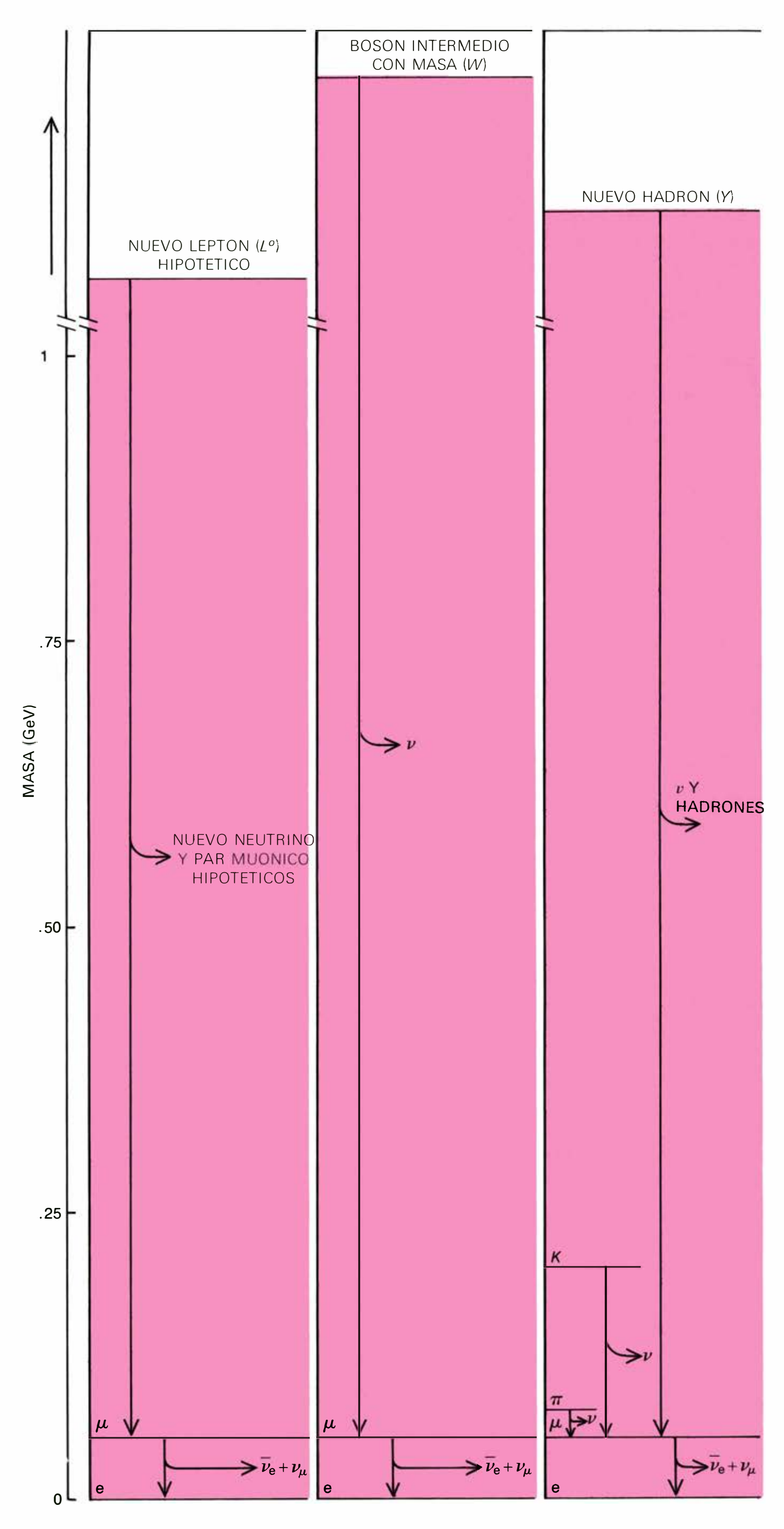
El método más prometedor para detectar nuevas partículas es el examen de los productos de ciertas interacciones en busca de alguna forma peculiar de desintegración. Como ya hemos dicho, los nuevos hadrones, los leptones pesados y las partículas W han de desintegrarse todas a través de la interacción débil. Los hadrones deberían realizar transiciones a hadrones ordinarios más ligeros, con la emisión de un muón o un electrón y un neutrino. Asimismo, se espera que entre los productos finales de desintegración xe los leptones pesados y las partículas W aparezcan un muón y un neutrino de tipo muónico (con valores opuestos del número cuántico L_μ). Por tanto, característica típica de la desintegración de estas partículas es la presencia de un neutrino por lo menos en el estado final.

Los neutrinos no se pueden detectar directamente; podrían pasar sin interacciones a través de cualquier aparato que pu xiéramos proyectar. Manifiestan, no obstante, su presencia por la conservación exacta de los números cuánticos L_e y L_μ. Los neutrinos finales escapan a toda detección, pero, como resultado de su existencia, la combinación de partículas restante muestra una aparente violación de la conservación de los números cuánticos leptónicos. Una vez aceptado que tal violación es imposible, se infiere la presencia de neutrinos.

Supongamos que una interacción débil se inicia mediante la colisión de un neutrino de tipo muónico con un neutrón. Si la reacción tiene lugar a través de una corriente cargada (lo cual es mucho más probable que a través de una corriente débil neutra), las leyes de conservación exigen que el estado final contenga un muón y, al menos, un barión, tal como el protón. La fórmula de la reacción sería: $\nu_\mu + n \rightarrow p + \mu^-$. Supongamos, sin embargo, que se creara una partícula W^+ en la interacción; la carga eléctrica podría conservarse sin necesidad de transformar el neutrón en un protón, siendo, entonces, la reacción $\nu_\mu + n \rightarrow n + W^+ + \mu^-$. No puede detectarse directamente la partícula W^+, pues se desintegraría muy rápidamente (en unos 10^{-18} segundos), pero los productos de su desintegración, un muón y un antineutrino, formarían parte del estado final. Así, la reacción total aparecería en la forma $\nu_\mu + n \rightarrow n + \mu^+ + \mu^- + \nu_\mu$, donde el neutrino de tipo muónico y el muón positivo del estado final constituyen los productos de la desintegración del W^+.

Resulta totalmente imposible detectar el neutrino saliente de la reacción anterior y, asimismo, costaría muchísimo detectar el neutrón saliente. Por lo tanto, el estado final aparecería como si estuviera formado por un muón positivo y un muón negativo, ambos fácilmente detectables. Un simple cálculo prueba que el número leptónico L_μ de las partículas detectadas en el estado final sería $-1+1$, es decir 0; en tanto que, para el estado inicial, L_μ era igual a $+1$. La aparente violación de la ley de conservación de L_μ revela que en la interacción había también un neutrino, y ello tiene la función de indicar que se ha creado una nueva partícula.

Si la nueva partícula hubiera sido un leptón pesado en lugar de un W^+, el estado final hubiera sido muy parecido. Si se hubiera tratado de un hadrón perteneciente a algún nuevo grupo, el estado final hubiera comprendido también dos muones, aunque no necesariamente muones de cargas eléctricas opuestas. En todos los casos, la existencia de un neutrino no detectado vendría indicada por una aparente violación de la conservación del número leptónico. Para distin-



LAS DESINTEGRACIONES DE LAS NUEVAS PARTICULAS conducen a la emisión de un muón y un neutrino. Para los hadrones *Y*, esta desintegración sería análoga a las desintegraciones del pión y del kaón, si bien las partículas *Y* lo harían con mayor rapidez. La masa de la partícula *Y* más ligera equivale a una energía entre 2 y 4 GeV (1 GeV = mil millones de electrón-volts); los bosones *W* tendrán una masa de 20 GeV por lo menos; las masas de los nuevos leptones son hipotéticas.

guir entre estas tres hipótesis sería nenecesario estudiar con mayor detenimiento las características de los dos muones.

A tales interacciones se las suele llamar *sucesos dimuónicos.* Nosotros hemos realizado una búsqueda pormenorizada de sucesos de este tipo en el Fermilab, en colaboración con diez físicos más de las Universidades de Harvard, y de Wisconsin y del Fermilab. Durante los dos últimos años hemos ido encontrando sucesos con las características previstas y posteriormente otro grupo de físicos del Instituto de Tecnología de California y del Fermilab confirmaron la observación de sucesos dimuónicos. El primer suceso dimuónico se observó en abril de 1973, poco después de la entrada en funcionamiento del acelerador del Fermilab.

En principio, las partículas *W* o los leptones pesados podrían generarse a través de muchas clases de interacciones, comprendidas las de un hadrón contra otro. Podrían obtenerse también miembros de nuevas familias de hadrones a partir de las colisiones entre hadrones ordinarios, pero sólo en pares partícula-antipartícula a causa de las rigurosas leyes de conservación que gobiernan las interacciones fuertes. Sin embargo, las colisiones de hadrones a altas energías son sucesos complejos en que otros muchos procesos compiten con los que se desean observar. Hubimos de adoptar, por lo tanto, una técnica distinta: bombardear nucleones con neutrinos de alta energía. En esos sucesos son posibles únicamente las interacciones débiles, y los resultados pueden interpretarse con relativa facilidad. Más aún, los nuevos hadrones pueden producirse de uno en uno, ya que las interacciones débiles pueden alterar los números cuánticos internos de los hadrones.

Lo primero que se necesita para aplicar esta técnica es un haz de neutrinos de alta energía. En el Fermilab se obtienen indirectamente a través de protones que se extraen del acelerador principal del laboratorio, un sincrotón de protones de 400 GeV. Los protones colisionan con un blanco metálico, generando un haz de hadrones secundarios, compuesto principalmente de piones y kaones. Estos mesones se desintegran al pasar a través de un tubo en el que se ha hecho el vacío, y que tiene unos 300 m de longitud. Como ya hemos visto, estos mesones se desintegran predominantemente en muones, neutrinos y antineutrinos. Al final de la región de desintegración existe un montículo de tierra de un

kilómetro de longitud, aproximadamente, que tiene por misión eliminar los muones y los posibles hadrones residuales. El haz que sale del montículo está formado, casi exclusivamente, por neutrinos y antineutrinos.

Los neutrinos deberán tener altas energías para poder crear partículas de gran masa. Es muy importante, asimismo, que el haz sea de elevada intensidad, esto es, que se halle formado por muchas partículas por segundo. La alta intensidad es esencial porque la probabilidad de que un neutrino determinado interaccione con un nucleón es muy pequeña. El haz del Fermilab alcanza una intensidad de unos 10^9 neutrinos por segundo; aun así, sólo un neutrino por minuto sufre una interacción. Todos los restantes atraviesan el aparato inalterados.

El detector que se empleó en los experimentos dimuónicos constaba de varios componentes dispuestos en secuencia a lo largo del eje del haz de neutrinos incidente. La primera sección estaba formada por cuatro cámaras que contenían unas 60 toneladas de líquido que centellea, es decir, que emite un destello luminoso, cuando una partícula con carga eléctrica lo atraviesa. La gran masa de líquido servía de blanco eficaz de nucleones para los neutrinos, y también absorbía la mayor parte de la energía de las interacciones. En cada tanque, unos tubos fotomultiplicadores detectaban los centelleos; la cantidad total de luz emitida daba una medida de la cantidad de energía depositada en el líquido. Entre los tanques de líquido había cámaras de chispa ópticas que permitían ver la cascada de partículas creadas en cada interacción.

Inmediatamente detrás del líquido de centelleo había una gran masa de hierro (también de unas 60 toneladas). El hierro servía de blanco adicional para las interacciones de los neutrinos, y, además, eliminaba algunas de las partículas no interesantes que se producían en el contador de centelleo. Más allá del apantallamiento de hierro había cuatro electroimanes de hierro, de forma cilíndrica, cada uno de los cuales tenía 3,6 metros de diámetro y una longitud de 1,2 metros.

Los imanes desviaban las partículas cargadas que pasaban por ellos, de modo que podía medirse el momento de dichas partículas. Las trayectorias de estas partículas se registraban por otras cámaras de chispas dispuestas entre los imanes.

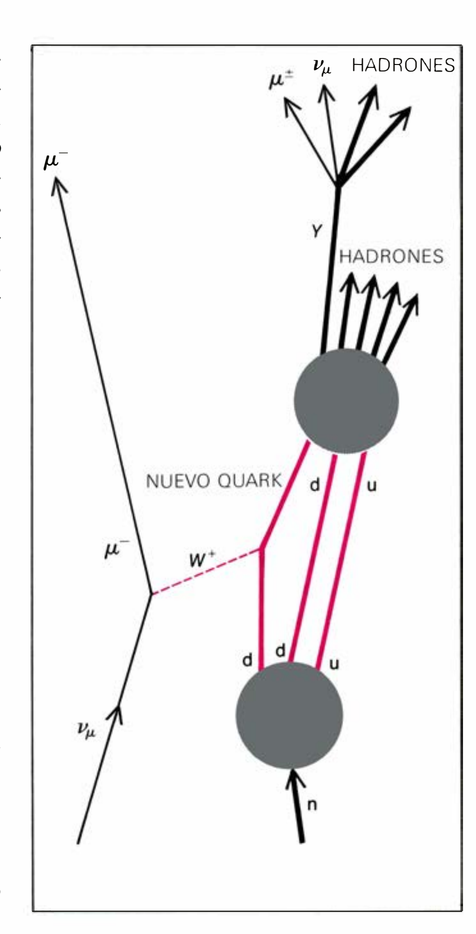
El detector se activaba a la llegada de una partícula cargada mediante un contador de disparo situado inmediatamente detrás del primero de los electroimanes. Tal partícula debía ser necesariamente un muón; ni el electrón ni un hadrón podrían pasar a través del blanco de hierro y del imán. Al activarse el detector, la señal eléctrica emitida por los contadores de centelleo quedaba registrada y se aplicaba un alto voltaje a las cámaras de chispas. Las trazas resultantes en las cámaras de chispas, que seguían los caminos ionizados por donde habían pasado las partículas cargadas, se registraban fotográficamente. Habíamos investigado la producción de dimuones buscando simplemente en las fotografías de cada suceso registrado la presencia de dos partículas cargadas que habían atravesado las barreras de hierro.

En varias series de experimentos, empleando haces con varias mezclas de neutrinos y antineutrinos, se registraron unas 80.000 activaciones del detector. Un total de 82 sucesos cumplían las condiciones que habíamos establecido para identificar sucesos dimuónicos auténticos. Para que un suceso se tuviera en cuenta, la primera exigencia era, por supuesto, que contuviera dos muones en el estado final. También era necesario que los muones se originaran en el mismo instante y en el mismo punto del espacio y que el momento de ambos pudiera medirse.

Para dar cuenta de los sucesos dimuónicos se puede apelar a distintas explicaciones que se fundan únicamente en procesos y en particulas convencionales. Una de las más obvias es la hipótesis de que un segundo muón surja de la desintegración de un pión o kaón cargado producido en la interacción del neutrino original. Sin embargo, a causa de las distintas densidades de los dos blancos, los muones provenientes de las desintegraciones de mesones se originarían con una frecuencia unas cuatro veces superior en el líquido de centelleo que en la barrera de hierro. En la práctica, se observa el mismo número de sucesos dimuónicos en el hierro y en el líquido, y esta igualdad excluye los mecanismos de desintegración de piones y kaones.

Otra posibilidad estriba en que la aparente violación de la ley de conservación del número cuántico leptónico L_μ no se deba a un neutrino, sino a un tercer muón todavía por detectar. Se inspeccionaron las fotografías de la cámara de burbujas en busca de sucesos con tres muones en el estado final. No se encontró ninguno.

Todas las explicaciones convencionales pueden refutarse o asignárseles una probabilidad muy baja. Quedan, pues,



LA HIPOTESIS DE LOS QUARKS puede explicar la creación de nuevos hadrones en las interacciones de los neutrinos. En el modelo quark, todos los bariones están formados por tres quarks; el neutrón está formado por los quarks *d, d* y *u*. A través de un bosón virtual W^+, el neutrino convierte uno de los quarks *d* en un quark de una nueva clase, que transporta un nuevo número cuántico. Los quarks interaccionan entonces para generar varios hadrones, entre los que se encuentra un (*Y*), por lo menos, con este número cuántico. Algunos identifican este nuevo número cuántico con la hipotética propiedad llamada "encanto".

tres explicaciones plausibles de los sucesos dimuónicos, las cuales tienen que dar razón de la producción y la desintegración de una de las tres partículas discutidas antes: un nuevo hadrón, un leptón pesado neutro o un bosón vectorial intermedio dotado de carga.

Para distinguir entre las tres posibilidades hay que tener en cuenta las propiedades de los pares de muones. Dos propiedades, en particular, son de la máxima importancia. En primer lugar, nuestras mediciones han demostrado que los dos muones no tienen en general momentos comparables. En los sucesos inducidos por neutrinos el muón negativo suele poseer un momento mayor que el muón positivo. En segundo lugar, los muones no tienen siempre cargas eléctricas opuestas; los sucesos en los que los dos muones tienen la misma carga cons-

tituyen aproximadamente un 10 % del total.

Ninguna de estas dos observaciones es compatible con la hipótesis de que los dimuones vengan generados por un leptón pesado o por una partícula *W*. La producción de muones con cargas iguales excluye concretamente la posibilidad de estos mecanismos. Puede pensarse que la interacción completa ocurra en el detector en dos estadios sucesivos: en el primero, el neutrino se convierte en otras partículas; en el segundo, el nucleón de retroceso se desintegra en otros hadrones. La observación de dos muones con la misma carga no puede explicarse entonces como resultado de la desintegración de ninguna partícula que se produzca en el primer estadio; y deben tenerse en cuenta, además, los hadrones del segundo estadio del proceso. Esta conclusión se ha visto confirmada al comparar los datos experimentales con cálculos teóricos llevados a cabo por Abraham Pais de la Rockefeller University y Sam B. Treiman de la Universidad de Princeton.

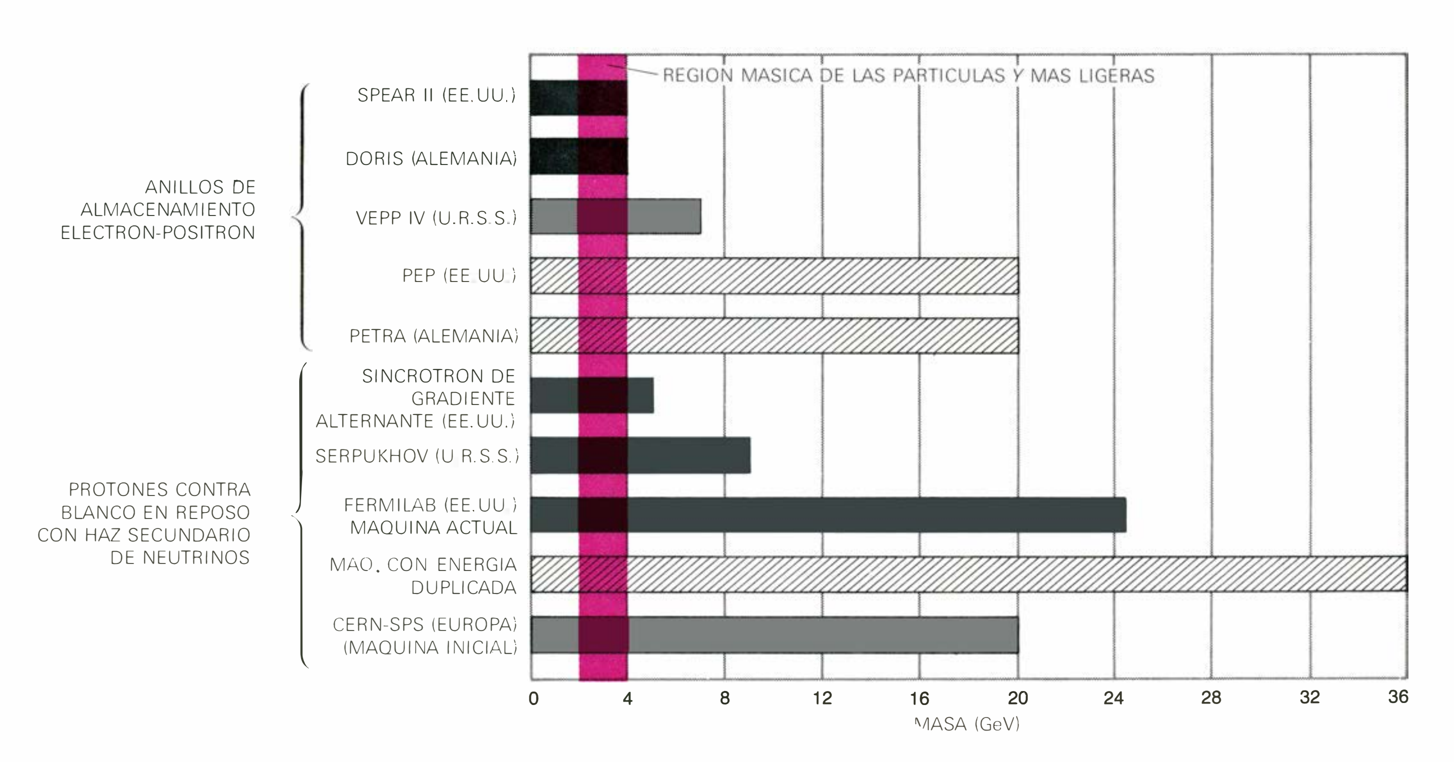
A partir de todo lo expuesto, hemos llegado a la conclusión de que la explicación más probable de los sucesos dimuónicos observados en el Fermilab es pensar en la creación y desintegración de un miembro, por lo menos, de una nueva familia de hadrones. A los miembros de esta nueva familia les hemos llamado partículas *Y*. Las partículas *Y* deben poseer necesariamente un nuevo número cuántico que se conserve en las interacciones fuertes y en las electromagnéticas, pero no en la interacción débil.

Deduciéndola de otras propiedades de los dimuones, hemos estimado que la masa de las partículas *Y* debe oscilar entre 2 y 4 GeV. Su vida media es sensiblemente inferior a los 10^{-8} segundos, y, con toda probabilidad, menor que 10^{-10} segundos. Al parecer, su desintegración origina leptones sólo en un 10 a 20 % de los casos; en las restantes desintegraciones se producen únicamente hadrones. Esto sugiere que las partículas *Y* se producen en sucesos con neutrinos de alta energía, con mucha más frecuencia de lo que a primera vista parecen indicar nuestros resultados.

Se precisarán nuevos experimentos antes de que la naturaleza y el significado de las partículas *Y* queden aclarados. Hemos supuesto la existencia de mesones y bariones *Y*, pero ello no está establecido de forma definitiva. Hay que medir con precisión las masas y vidas medias, y descubrir sus modos de desintegración. A pesar de tales lagunas, las partículas *Y* representan, hoy por hoy, la prueba más directa y substancial de la existencia de hadrones con un nuevo número cuántico. Las partículas *Y* parecen mostrar directamente ese número cuántico, en tanto que las partículas *J* o psi reflejan sólo de forma indirecta la existencia de un nuevo número cuántico. No se sabe si es el mismo número cuántico lo que distingue ambos estados.

Tanto la partícula *Y* como las partículas *J* o psi pueden interpretarse a la luz de un complejo modelo teórico de hadrones introducido en 1963 por Murray Gell-Mann y George Zweig del Instituto de Tecnología de California. En ese modelo, los hadrones se consideran estructuras compuestas, formadas a partir de constituyentes más simples llamados quarks. Los bariones constan de tres quarks y, los mesones, de un quark y un antiquark. En la formulación original del modelo había tres clases de quarks, denominados *u, d* y *s*. El quark *s* tiene extrañeza 1, y todas las partículas extrañas poseen, al menos, un quark *s* o un antiquark $\bar{s}$.

En el modelo de quark, los desdoblamientos másicos entre hadrones son un puro reflejo de los desdoblamientos másicos de los quarks. El hecho de que el quark *s* sea considerablemente más pe-



DIEZ ACELERADORES DE PARTICULAS, capaces de explorar la región másica de las partículas *Y*, entrarán en funcionamiento en la próxima década. Cinco de ellos están actualmente en funcionamiento *(en negro)*; dos más, en construcción *(en gris)*; los tres restantes se hallan en fase de proyecto *(barras inclinadas)*. En los anillos de almacenamiento electrón-positrón, las partículas chocan con velocidades iguales y opuestas; en los otros aceleradores, los neutrinos chocan con los nucleones en un blanco fijo. Las masas no representan la energía requerida para la aceleración de las partículas, sino la masa máxima de una partícula que puede crearse por las interacciones de las partículas aceleradas. En el Fermilab, la energía del sincrotrón de protones doblará tras la instalación de imanes superconductores, lo que incrementa la energía de las interacciones de neutrinos en un 50 %. Algunos podrán crear leptones pesados, bosones *W* y nuevas familias de hadrones.

sado que los otros quarks explica que los hadrones con extrañeza pesen más que los equivalentes no extraños.

Muy recientemente se ha avanzado la hipótesis de que pudiera haber un cuarto quark, dotado de un nuevo número cuántico que se conserva en las interacciones fuertes y electromagnéticas. Sheldon Lee Glashow de Harvard y James D. Bjorken del Centro del acelerador lineal de Stanford, propusieron que el nuevo número cuántico se llamara "encanto" *(Charm),* y el cuarto quark se le denomina ahora quark *c* o encantado. Si el quark encantado existiera, entonces debe existir una nueva clase de hadrones, que se desintegrarían en hadrones ordinarios a través únicamente de la interacción débil.

Más aún, como el quark encantado tendría una masa mucho mayor que la del quark *s*, los desdoblamientos másicos de los nuevos hadrones serían mayores que los que se observan entre los hadrones ordinarios.

Si las partículas *Y,* con sus nuevos números cuánticos, hubieran de incorporarse en el modelo de quark, se exigirá, por lo menos, un nuevo quark. Parece razonable identificar el cuarto quark con el encantado propuesto por Gashow y Bjorken; identidad que sólo podrá demostrarse con nuevos experimentos. La hipótesis del encanto viene avalada por vigorosos razonamientos teóricos independientes de los descubrimientos recientes; más aún, puede explicar los estados *J* o psi y los *Y*. Los estados *Y* se supone que están formados de quarks *c* en combinación con los de las otras clases, mientras que la *J* o psi sería un estado ligado de un quark *c* y un antiquark $\bar{c}$.

Todo parece, pues, que teoría y experimentación convergen sobre una nueva familia de hadrones, de forma análoga a lo que sucedió hace aproximadamente unos 25 años.

Resulta paradójico que la observación de los sucesos dimuónicos inducidos por neutrinos y antineutrinos conduzca al descubrimiento de un nuevo número cuántico para los hadrones. Durante muchos años, los sucesos dimuónicos fueron objeto de minuciosa investigación como prueba de la producción de bosones vectoriales intermedios con masa.

Ahora que, por fin, hemos dado con la prueba, se la asigna un significado totalmente distinto.

¿Qué perspectivas hay de descubrir nuevas partículas? Una conclusión que ya parece justificada es que la masa de las familias *W* es mucho mayor que la masa de los nuevos hadrones. Hemos determinado por vía experimental que la masa de los bosones *W* cargados debe superar los 20 GeV. De más difícil pronóstico son la masa y las restantes propiedades de los leptones pesados. Pueden haber varios tipos, unos más fáciles de crear y detectar que otros. Por ejemplo, puede haber leptones pesados con una masa menor que la que poseen los hadrones *Y*.

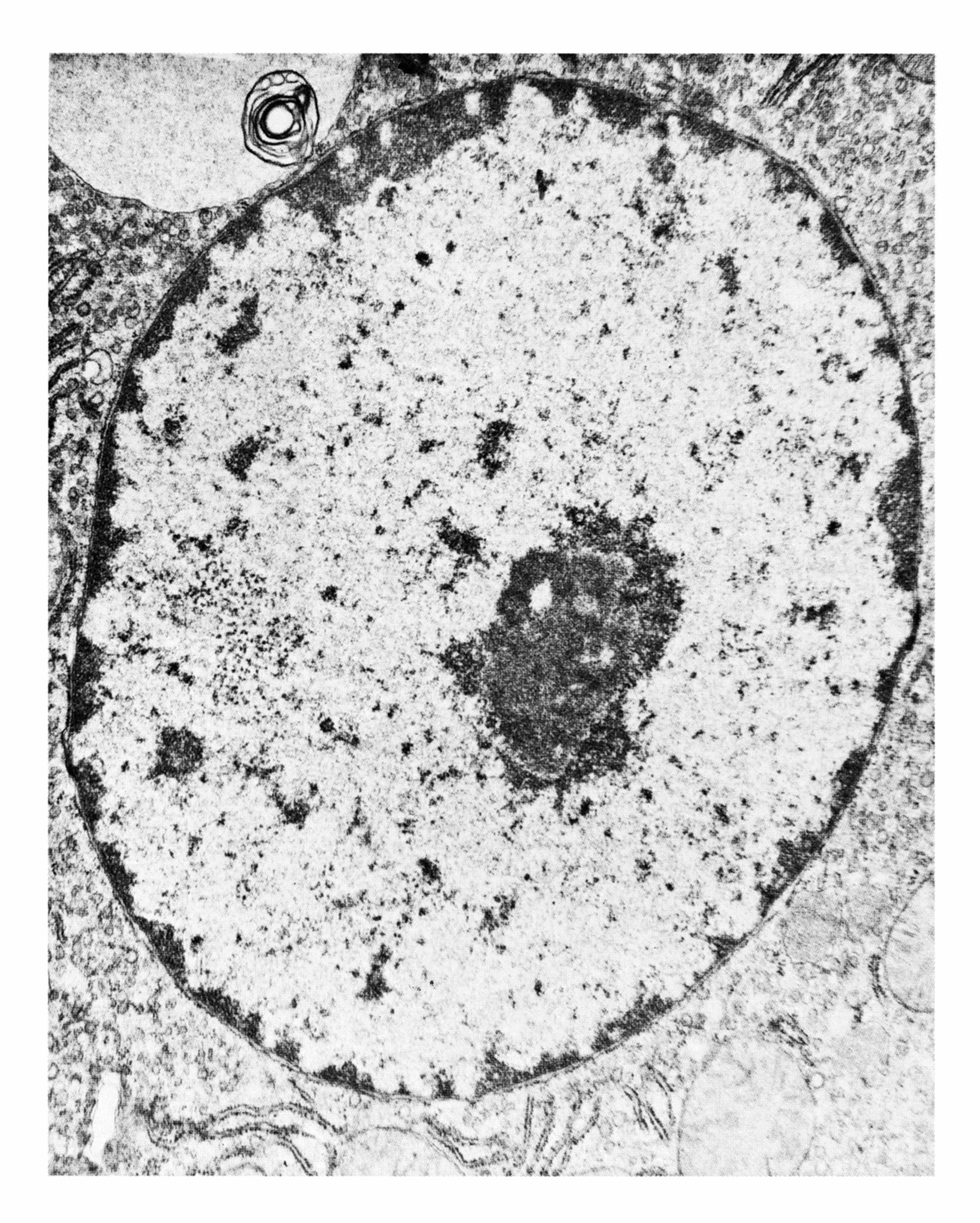
Las partículas *Y* podrían entonces desintegrarse, a través de estos leptones, en leptones ordinarios y una nueva clase de neutrino; el estadio final del proceso se manifestaría como un suceso dimuónico.

En este caso, se hubieran producido ambas clases de partículas nuevas en un único experimento. Por ahora, disponemos de pocos datos para excluir esta posibilidad.

Una consecuencia inmediata que podría inferirse de la existencia de mesones *Y,* con masas entre 2 y 4 GeV, es que los estados mesónicos con masa menor en cada uno de los grandes grupos de partículas mantienen una distribución simple: sus masas se hallan espaciadas en forma más o menos regular. Parece que tenga que existir una relación simple entre la masa del mesón pi, la del mesón *K* y la del mesón *Y* de masa inferior; si se confirma esta relación, pudiera pensarse que exista un nuevo mesón, con otro nuevo número cuántico, en la región de masas comprendida entre 8 y 15 Gigaelectronvolts.

Es probable que la búsqueda de tales mesones constituya, en los próximos años, el objetivo de los físicos experimentales, en particular de aquellos que experimenten con neutrinos de alta energía.

El afán descubridor de nuevos estados mesónicos y otras partículas vendrá facilitado por aceleradores nuevos y otras máquinas, comprendidas algunas que ya han entrado en funcionamiento y otras que se están construyendo. En el Fermilab, se está modificando el principal sincrotón de protones con la incorporación de imanes superconductores que doblarán su energía. Con nuevas máquinas de esta índole se podrán explorar regiones de masa mayor, y, con métodos indirectos para la detección de partículas, se podrán determinar las relaciones y las propiedades de las nuevas familias. A pesar de la asombrosa proliferación de partículas y de números cuánticos hay esperanzas de que el modelo de la naturaleza resultante de estas investigaciones sea sencillo y atractivo.



NUCLEO DE UNA CELULA DE HIGADO DE RATON, aumentado 22000 veces en esta microfotografía electrónica de David E. Comings y Tadashi Okada del City of Hope Medical Center. El núcleo, rodeado por su fina membrana doble, está repleto de cromatina, substancia encargada de las funciones genéticas y de control de la célula. En la cromatina, el ADN o material genético forma complejos con las proteínas cuya misión parece ser regular la actividad de los genes. Las manchas oscuras que están junto a la membrana y en otras partes de la célula son de cromatina "condensada", que no sintetiza ARN; la substancia más clara que llena gran parte del núcleo es cromatina "extendida", encargada de la síntesis activa de ARN. A la izquierda del centro del núcleo, se distinguen gránulos de ADN, que forman complejos con las proteínas. La estructura mayor, también oscura, situada en la zona central, es el nucleolo, que contiene cromatina especializada.

Las proteínas cromosómicas y la regulación de los genes

Se está empezando a comprender la función de las proteínas en los núcleos celulares de los organismos superiores. Al parecer, las proteínas histónicas inactivan a los genes y las no histónicas los activan selectivamente

Gary S. Stein, Janet Swinehart Stein y Lewis J. Kleinsmith

Como es bien sabido, los genes, que transmiten la información hereditaria de generación en generación y dirigen el funcionamiento de la célula viva, están constituidos por ADN. Sin embargo, en las células de los organismos superiores, los genes se hallan dispuestos en cromosomas, y éstos son nucleoproteínas, a saber: complejos de ácidos nucleicos y proteínas. Si el ADN constituye el material genético, ¿qué función corresponde a las proteínas cromosómicas? Al parecer, desempeñan un papel importante en el mantenimiento de la estructura del material genético y en la regulación de la actividad genética, es decir, en la determinación de qué genes se activan en una célula dada y cuándo lo hacen. El control de los genes es de importancia suma en procesos fundamentales, como la diferenciación, el desarrollo embrionario y la acción hormonal, y en procesos anormales, verbigracia, el cáncer, las enfermedades metabólicas y las anomalías congénitas, que están relacionadas con la expresión de la información genética. En este artículo nos dedicaremos principalmente a estudiar la actividad reguladora de las proteínas cromosómicas, un campo de investigación que ha comenzado a proporcionar resultados significativos en los últimos años.

Regulación génica

La información genética de una célula está codificada en las secuencias de nucleótidos que constituyen la molécula de ADN. Para utilizar esta información, la célula transcribe las secuencias de nucleótidos en cadenas complementarias de ARN que se traducen en cadenas de aminoácidos, las cuales forman, a su vez, las proteínas. El control de la actividad génica reside en la selectividad con que se realiza la transcripción en el tiempo y en el espacio. En el desarrollo embrionario, por ejemplo, un solo huevo fecundado, que contiene todos los genes del organismo, prolifera hasta dar lugar a una gran variedad de células diferenciadas, que se especializan en funciones características. A lo largo del desarrollo se activan e inactivan cientos de miles de genes. Además, en cada tipo final de células se transcriben, o se expresan, sólo ciertas combinaciones de genes, de acuerdo con la función específica de la célula. En los glóbulos rojos, por ejemplo, puede tratarse de la actividad del gen que codifica para la hemoglobina; en las células musculares, el gen para la mioglobina; en los fibroblastos, el gen para el colágeno, que es la proteína del tejido conjuntivo.

A pesar de lo cual, por lo que hasta ahora se sabe, el ADN de todos estos tipos de células especializadas es el mismo. La constancia del ADN en todas las células de un organismo dado se estableció por primera vez en 1948 y se ha confirmado de diversas maneras, siendo la más atrayente la experimentación llevada a cabo por J. B. Gurdon, de la Universidad de Oxford. Gurdon transplantó núcleos de células intestinales diferenciadas de rana en huevos de rana cuyos núcleos celulares se habían extirpado. Los huevos se desarrollaron normalmente.

Tales resultados demostraron que los núcleos de las células intestinales, y todas las demás células diferenciadas de la rana, contenían la información genética completa de ese animal, aun cuando sólo se exprese una porción limitada de esta información.

Dentro, incluso, de una célula diferenciada dada, se expresa cierta información en un momento determinado y permanece sin expresar en otro. La expresión puede modificarse cuando las células que están en reposo se activan para dividirse y cuando son estimuladas por la acción de hormonas específicas; en ambos casos, se operan cambios complejos en el metabolismo celular al ocurrir variaciones en el programa genético de la célula. En general, en cualquier momento se expresa menos del diez por ciento de la información genética total. Los mecanismos reguladores específicos activan o inactivan para la transcripción ciertas regiones apropiadas del genoma, que es la dotación total de genes, conforme van cambiando las necesidades de la célula.

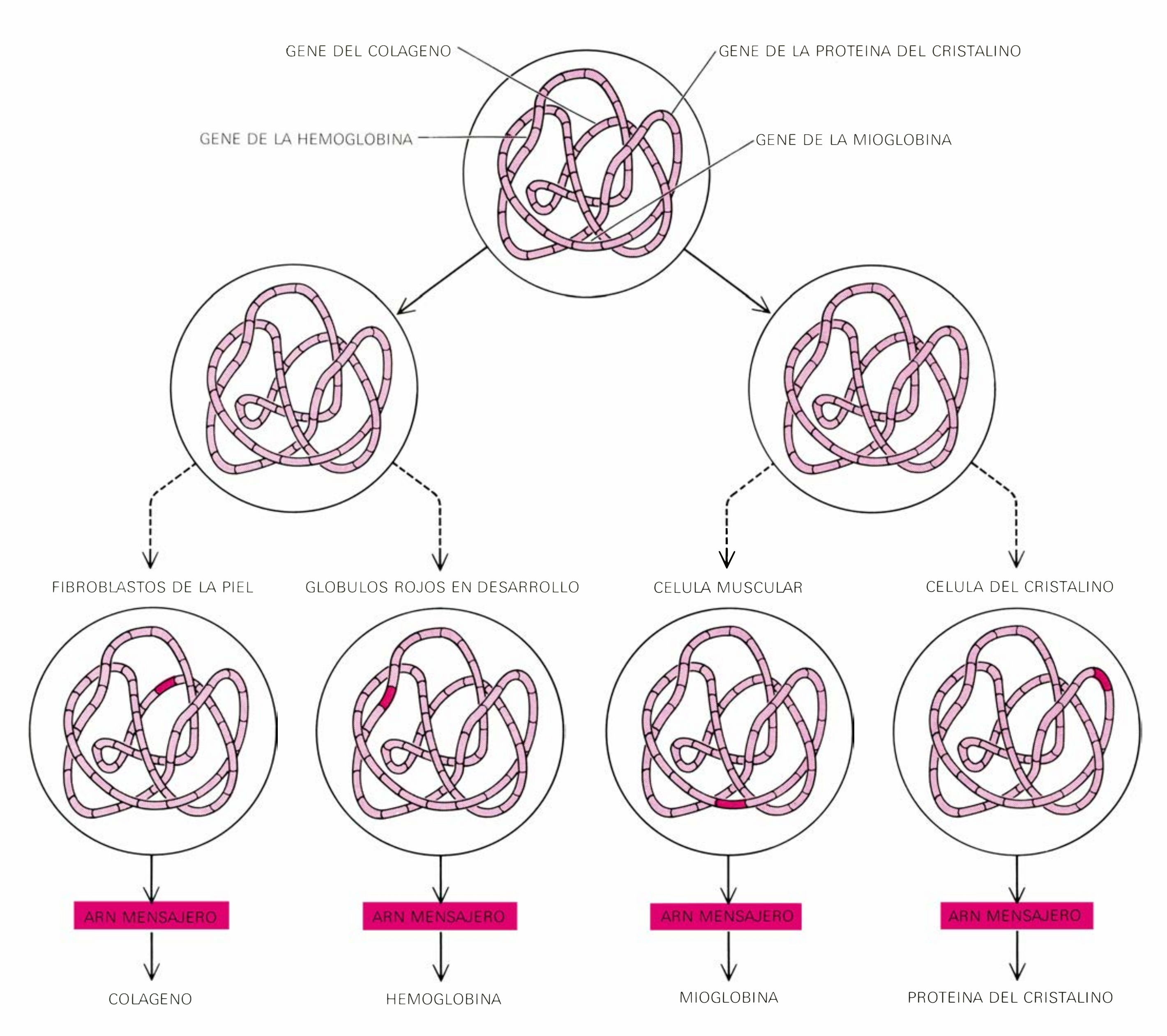
En las bacterias, que no tienen núcleo y cuyo genoma es sensiblemente menos complejo que el de los organismos superiores, se ha llevado a cabo un progreso sustancial en favor de la dilucidación de los mecanismos reguladores. En el bacilo del colon, *Escherichia coli*, se ha demostrado que la transcripción de secuencias genéticas particulares en ARN está regulada por la unión y desunión de reguladores específicos en lugares apropiados del ADN. El material genético es más complejo estructural y funcionalmente en las células de los vegetales y de los animales superiores. El ADN se contiene en un núcleo al que rodea una membrana. Forma allí un conglomerado con proteínas y con una pequeña cantidad de ARN, que constituye una sustancia compleja llamada cromatina o, en su forma más organizada, los cromosomas. Hasta ahora no han podido definirse proteínas reguladoras específicas en tales células nucleadas, o eucariotas, pero los datos

indican que las proteínas cromosómicas constituyen los elementos de regulación.

Las histonas

Desde un punto de vista histórico, tales proteínas se han dividido en dos grupos: las histonas y las proteínas no histónicas. Las características químicas de las histonas nos dan la mejor definición de las mismas: son proteínas con carga positiva, ricas en los aminoácidos básicos arginina y lisina, carentes en absoluto del aminoácido triptófano. Dada su naturaleza básica, pueden extraerse fácilmente con ácido clorhídrico o sulfúrico diluido. Pueden catalogarse en cinco clases principales, muchas de las cuales están presentes en todas las células eucariotas. Las histonas se descubrieron a fines del siglo XIX, pero hasta 1943 Edgar y Ellen Stedman, trabajando a la sazón en la Universidad de Edinburgo, no fijaron en ellas la atención como posibles moléculas reguladoras. Los datos de los Stedman sugerían que los tejidos en fase de crecimiento activo contenían menos histona que los tejidos en otras fases; llegaron, por tanto, a la conclusión de que las histonas funcionaban como represores, o inhibidores, de la actividad biológica. Postularon correctamente una función inhibitoria para las histonas, pero sus conclusiones se basaban en mediciones del contenido en histona, que se ha demostrado después que eran inexactas; hoy sabemos que el contenido en histona es idéntico para las células activas e inactivas.

Los primeros estudios bioquímicos decisivos sobre los efectos de las histonas en la función del ADN se emprendieron por los primeros años de la década de los 60, después del desarrollo de los sistemas acelulares para la síntesis de ARN. En estos sistemas, el ADN o la cromatina pueden servir como moldes para la transcripción del ARN en el tubo de ensayo. Otros componentes de

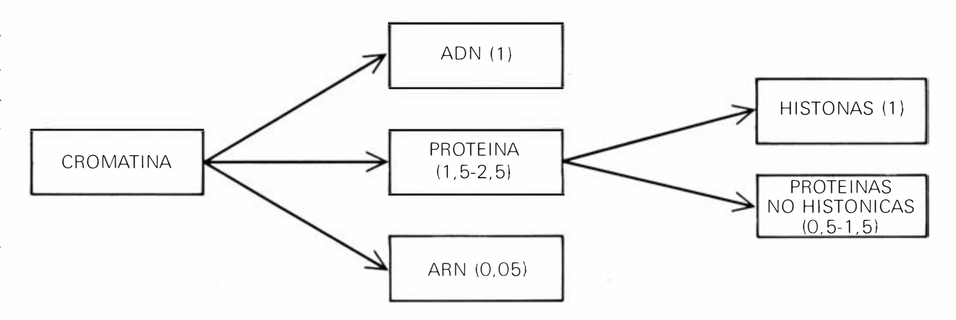


LA REGULACION GENICA en el curso del desarrollo. El huevo fecundado *(arriba)* contiene toda la información genética del organismo, codificada en el ADN; están los genes para las cuatro proteínas especiales que se indican y miles más. El huevo se divide, las células proliferan y se diferenciarán *(flechas de trazo discontinuo)* en células especializadas que han de constituir los distintos tejidos. La célula diferenciada sigue conteniendo todos los genes del organismo, pero éstos están regulados de tal manera que sólo se activan los adecuados para cumplir las funciones específicas. Es decir, se transcriben dando ácido ribonucleico (ARN) mensajero y, de aquí, se traducen en proteínas *(abajo)*.

la mezcla reaccionante son los precursores de las subunidades del ARN y la ARN polimerasa, la enzima que cataliza la polimerización de las subunidades para dar las cadenas de ARN. En 1962, Ru-chih Huang y James Bonner, del Instituto de Tecnología de California, demostraron que la adición de histonas a tal sistema acelular inhibía la síntesis de ARN. (La inhibición máxima de la síntesis de ARN se observó con una relación histona/ADN de 1:1, que coincide casi con la relación en que la histona y el ADN se encuentran normalmente en el núcleo.) La inhibición se anulaba mediante la adición de más ADN, lo cual demostraba que las histonas bloqueaban la transcripción uniéndose al ADN más que inhibiendo a la ARN polimerasa. Casi al mismo tiempo, Vincent G. Allfrey y Alfred E. Mirsky de la Rockefeller University demostraron que la eliminación selectiva de histonas de núcleos aislados aumentaba la tasa de síntesis de ARN. Ambos tipos de experimentos llevaron a la conclusión de que las histonas funcionan normalmente inhibiendo la capacidad para que los genes se transcriban en el ARN.

Muy pronto pareció evidente que las histonas se hallan implicadas también en la determinación de las propiedades estructurales de la cromatina. En particular, la adición de histonas aumenta la estabilidad de la doble hélice del ADN. Este hecho proporciona una pista sobre la función biológica de las histonas, dado que un estrechamiento de la hélice podría afectar a la disponibilidad, o exposición, de las secuencias de nucleótidos que se transcriben en el ARN.

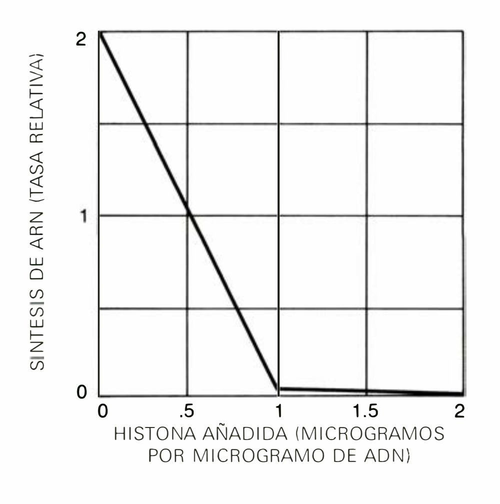
Otro aspecto del metabolismo de las histonas que evoca su función biológica reside en la relación entre la síntesis de histonas y la replicación del ADN. Antes de que una célula se divida para dar otras células hijas, debe duplicarse todo su ADN; de manera que cada célula hija reciba una copia completa de la información genética. La síntesis de ADN está restringida a un período definido del ciclo vital de la célula, la fase *S*; y está estrechamente acoplada a la síntesis de histonas. Thaddeus Borum, Elliot Robbins, y Matthew D. Scharff, del Albert Einstein College of Medicine, y Gerald C. Mueller, de la Universidad de Wisconsin, demostraron que las histonas y el ADN se sintetizaban simultáneamente y que la inhibición de la replicación del ADN se hallaba acompañada de paro repentino en la síntesis de histona. El estrecho acoplamiento entre síntesis de ADN y síntesis de histona está confirmado, además, por experimentos sobre el ARN; éstos muestran que los ARN mensajeros que codifican para las histonas se encuentran asociados únicamente a los orgánulos celulares encargados de la síntesis de proteínas y que se traducen dando proteínas al tiempo que se sintetiza el ADN, lo cual significa que, si se tienen que transcribir sólo regiones limitadas del genoma en una célula dada en un momento determinado, las histonas tienen que estar inmediatamente disponibles para reprimir los segmentos de ADN recién replicados, que habrán de quedar sin expresión.

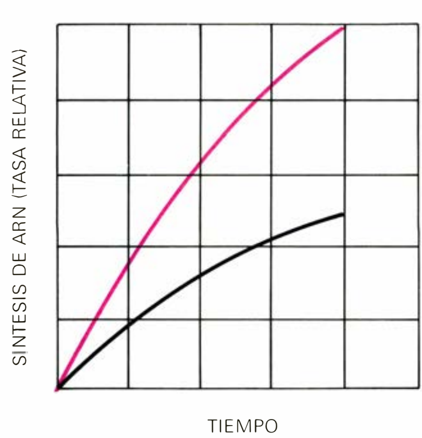


LA CROMATINA es un material complejo presente en los núcleos de las células de los organismos superiores, formado por ADN, proteínas y ARN. Hay dos tipos de proteínas: las histonas y las proteínas no histónicas. Se indican las cantidades relativas de cada componente. Hay, casi, la misma cantidad de histonas que de ADN, pero la relación varía con las proteínas no histónicas.

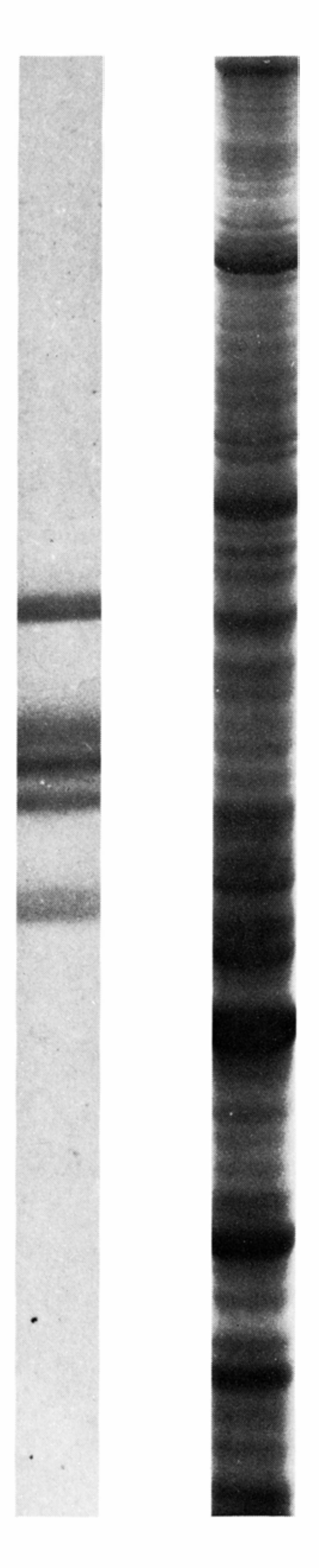
Aparte de la síntesis de nuevas moléculas de histona, la interacción de éstas con el ADN puede alterarse mediante la modificación de las histonas ya existentes. Hay reacciones enzimáticas que modifican la estructura de las histonas por la adición de grupos acetato, metilo o fosfato. No se conoce todavía el significado exacto de estas reacciones, pero Allfrey y colaboradores han demostrado que la acetilación de la histona está relacionada con el nivel general de síntesis de ARN; G. Roger Chalkley, de la Universidad de Iowa, ha observado que la fosforilización de ciertas histonas se halla en función de la velocidad de división celular. Tales reacciones pueden ser importantes en el control de la estructura y actividad globales del genoma.

En resumen, queda claro que las histonas están implicadas en el mantenimiento de la estructura de la cromatina así como en la represión de la síntesis de ARN dependiente del ADN. No obstante, determinadas series de pruebas sugieren que las histonas muestran una uniformidad, y una consecuente falta de especificidad, que obstaculizan su capacidad para reconocer e influir sobre genes particulares. Cantidades similares de histona están presentes en tejidos activos e inactivos, y su cantidad por célula permanece constante conforme varía el metabolismo celular. Se encuentran también presentes cantidades similares de histona en las regiones activas e inactivas del genoma de una célula dada, es decir, en la cromatina "extendida", que está sintetizando activamente ARN; y en





EL EFECTO INHIBIDOR de las histonas en la actividad génica se demostró por primera vez en los años sesenta. Ru-chih Huang y James Bonner, del Instituto de Tecnología de California, observaron *(izquierda)* que éstas inhiben la síntesis de ARN en sistemas acelulares. Allfrey y Mirsky, de la Rockefeller University, demostraron *(derecha)* que la tasa de síntesis de ARN era mayor en los núcleos de los que se habían eliminado las histonas *(color)* que en los núcleos de control *(negro)*.



LAS HISTONAS Y LAS PROTEINAS NO HISTONICAS, aisladas de la cromatina, se sometieron a electroforesis en tubos de gel de poliacrilamida, para su separación, ya que se desplazan a distinta velocidad en un campo eléctrico. Las histonas *(a la izquierda)* se distribuyen en cinco grupos definidos; la fracción de proteínas no histónicas *(a la derecha)* es mucho más heterogénea. Al analizar sus aminoácidos componentes se ha observado que las histonas abundan en aminoácidos básicos, en tanto que las proteínas no histónicas tienen un contenido muy alto de aminoácidos ácidos.

la cromatina "condensada", que no lo está sintetizando. En los núcleos de determinadas células de larvas de insectos, las regiones del genoma que están en el proceso de síntesis de ARN se forman conglomerados *(puffs)* relativamente voluminosos. Hewson Swift, de la Universidad de Chicago, midió la relación histona/ADN en las regiones inactivas y en los conglomerados y encontró que eran idénticas. Quizás la prueba más concluyente de la no especificidad de las histonas sea la dada por la comparación entre las secuencias de aminoácidos de una misma fracción de histona en dos tipos muy diferentes de células, las de plantas jóvenes de guisante y las de timo de ternera; resultaba que sólo diferían en dos aminoácidos. Parece claro, pues, que las histonas sean moléculas reguladoras, comprometidas en el control de la transcripción genética, pero deben serlo en un sentido no específico.

Las proteínas no histónicas

La investigación sobre moléculas capaces de regular genes específicos ha puesto la atención recientemente en las proteínas no histónicas de los cromosomas, que se definen simplemente como aquellas proteínas distintas de las histonas que están asociadas a la cromatina. Igual que las histonas, se sintetizan fuera del núcleo, en el citoplasma. Inmediatamente después de su síntesis las histonas se transportan al núcleo y se asocian al ADN. Las investigaciones llevadas a cabo en nuestros laboratorios de la Universidad de Florida (College of Medicine) y en la Universidad de Michigan han demostrado que, para las proteínas no histónicas, el panorama es algo más complejo. Algunas proteínas no histónicas, inmediatamente después de su síntesis, pasan a asociarse al ADN, pero otras tardan un tiempo variable en aparecer como componentes de la cromatina. En tanto que grupo, las proteínas no histónicas muestran una tasa de recambio superior a las histonas, con una oscilación considerable en el tiempo de vida de cada una de las fracciones no histónicas; unas perviven escasos minutos mientras que otras tienen por lo menos la estabilidad del ADN celular y de las histonas. No está claro si, al abandonar su asociación con el ADN, las proteínas no histónicas se degradan o entran a formar parte de una reserva de proteínas en el núcleo o en el citoplasma. Lo que sí queda patente es que, así como las histonas y el ADN parecen ser componentes permanentes del genoma, las proteínas cromosómicas no histónicas están, al menos parcialmente, en estado de flujo dinámico.

En contraposición a las histonas, las proteínas no histónicas manifiestan una gran heterogeneidad desde el punto de vista estructural y funcional *(véase la figura de la izquierda)*. Su gama de pesos moleculares oscila desde menos de 10.000 daltons hasta más de 150.000. Es difícil determinar exactamente cuántos tipos de proteínas no histónicas se hallan asociados al genoma, pero hay varias clases de pruebas que sugieren que es un número muy elevado. Muestra de su diversidad funcional es la variedad de sistemas enzimáticos complejos encontrados entre ellas. Hay enzimas, tales como las polimerasas, implicados en la síntesis y reparación del ARN y ADN, enzimas implicados en la síntesis y degradación de las proteínas y enzimas que modifican los ácidos nucleicos y las proteínas añadiendo o eliminando grupos acetato, metilo o fosfato. Además de estos componentes enzimáticos, se cree que la fracción no histónica contiene también proteínas con una misión estructural y reguladora en la cromatina. En adelante, sin embargo, habremos de ceñirnos a estudiar las argumentaciones principales que corroboran la idea de que algunas proteínas no histónicas, por lo menos, se hallan comprometidas en la regulación de la actividad de genes específicos.

Hay notables diferencias entre los tipos de proteínas no histónicas encontradas en células de especies distintas. Dado que la información genética codificada por el ADN es única para cada especie, la especificidad de especie observada en las proteínas no histónicas está de acuerdo con la función reguladora. Además, en el mismo organismo, las proteínas no histónicas varían según los tejidos. Aunque todas las células de un organismo contienen el mismo ADN, la variación de la dotación de proteínas no histónicas en los distintos tipos celulares es coherente con la utilización selectiva de la información genética por parte de tales tipos celulares.

Si las proteínas no histónicas están encargadas de la regulación génica específica, el espectro de estas proteínas en una célula debería cambiar cuando lo hace la actividad génica. Esta hipótesis se está sometiendo a comprobación en gran número de sistemas experimentales donde se sabe que hay variaciones notorias en la expresión génica. En casi todos los casos se han observado las modificaciones previstas en las proteínas no histónicas.

Variaciones en la síntesis

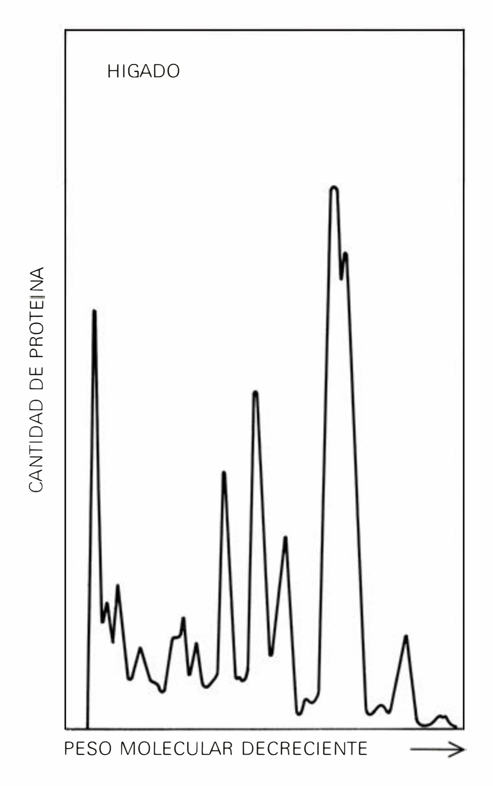
La expresión génica varía, por ejemplo, a lo largo del ciclo celular. Las diferencias más acusadas en la transcripción se observan entre la fase *S* (período de síntesis del ADN) y la mitosis (proceso de división celular). En nuestros laboratorios y en otros, se han examinado las proteínas cromosómicas no histónicas asociadas al genoma, durante cada fase del ciclo celular, en lo que respecta a su composición y metabolismo; se han observado diferencias de parecido alcance a los ocurridos en la transcripción.

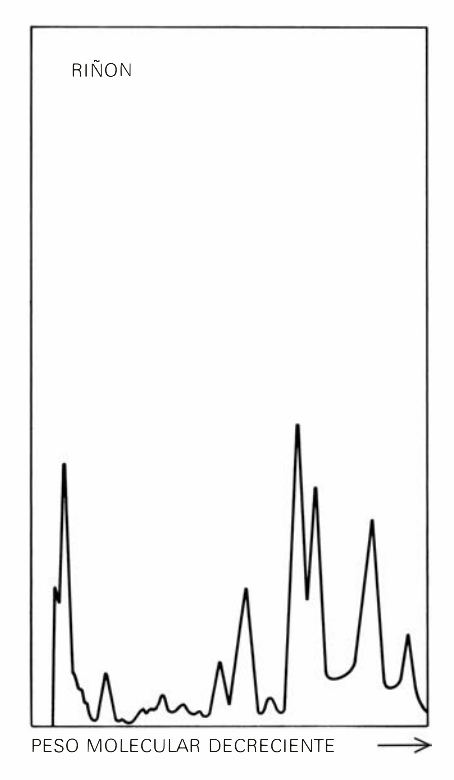
Otro ejemplo más: las hormonas esteroides; se sabe que los estrógenos y la hidrocortisona ejercen sus efectos en las células diana o receptoras uniéndose al núcleo y alterando, finalmente, el modelo de transcripción génica. Se ha demostrado que varias hormonas esteroides inducen cambios en la composición y en el metabolismo de las proteínas cromosómicas no histónicas.

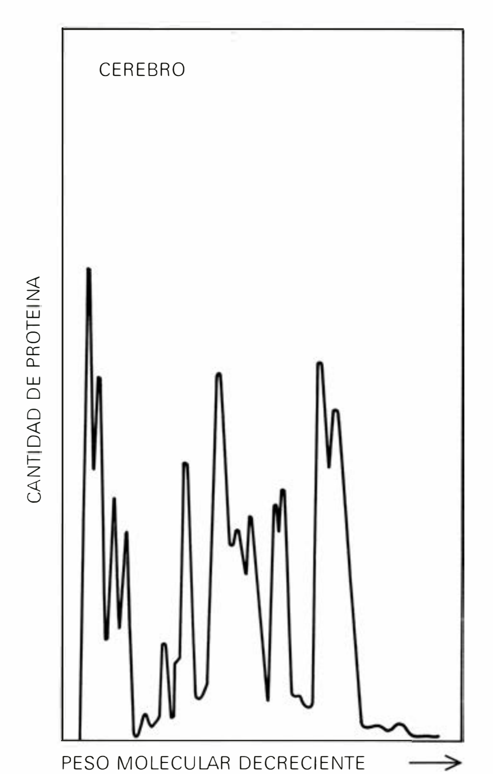
Los cultivos "in vitro" de células del neuroblastoma de ratón (cáncer de las células nerviosas) proporcionan un modelo interesante para estudiar la diferenciación. Se las puede inducir fácilmente a que se diferencien en células con las propiedades estructurales, bioquímicas y eléctricas características de las células nerviosas. Una de las ventajas peculiares de este sistema es que la diferenciación inducida sea completamente reversible, Michelle Zornetzer observó, en nuestro laboratorio de Florida, que en el estado diferenciado existe una disminución selectiva de la síntesis de proteínas no histónicas de elevado peso molecular y que este fenómeno revierte cuando lo hace el estado de diferenciación *(véase la ilustración de la página 38).* Demostró también que la diferenciación de tales células depende de las variaciones presentadas en la síntesis de ARN. Hay, pues, una estrecha relación entre los tipos de proteínas no histónicas asociadas al genoma y las alteraciones de la expresión génica que acompaña a la diferenciación.

La cromatina puede separarse en dos fracciones que divergen en la estructura y en la actividad genética: la fracción de cromatina condensada, estructuralmente compacta y relativamente inactiva en la síntesis de ARN, y la fracción de cromatina extendida, que sintetiza activamente el ARN. Como hemos dicho, la relación histona/ADN es la misma en ambos tipos de cromatina. Sin embargo, la cantidad relativa de proteína no histónica es considerablemente mayor en la forma extendida y activa de la cromatina. Una vez más la presencia de proteínas no histónicas se halla en función de un estado estructural de la cromatina, caracterizado por la síntesis activa de ARN.

A las modificaciones en las propiedades estructurales y bioquímicas de las células infectadas y transformadas por virus cancerígenos, acompañan también los cambios en las proteínas no histónicas. Las células transformadas ma-



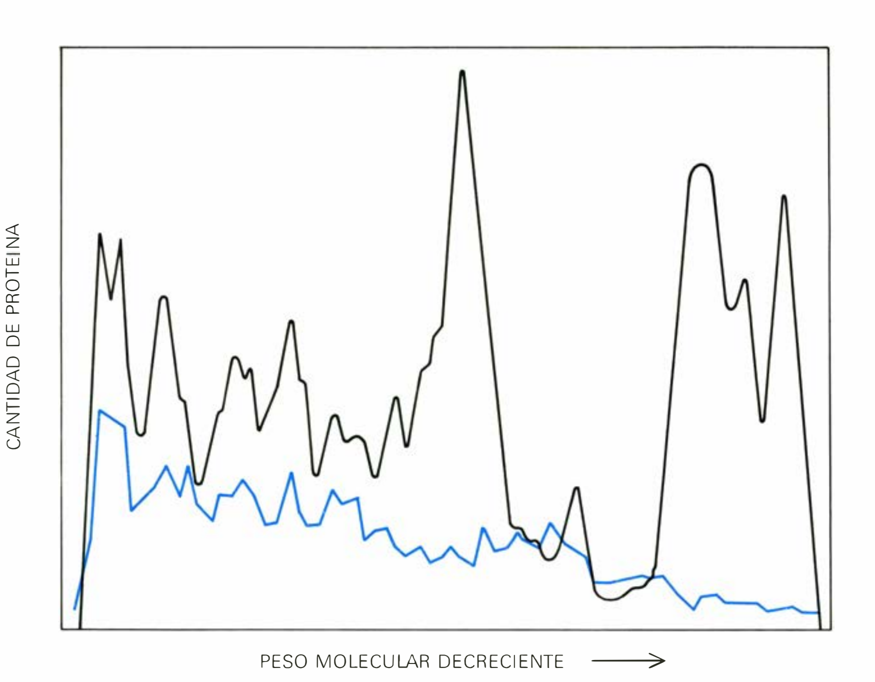


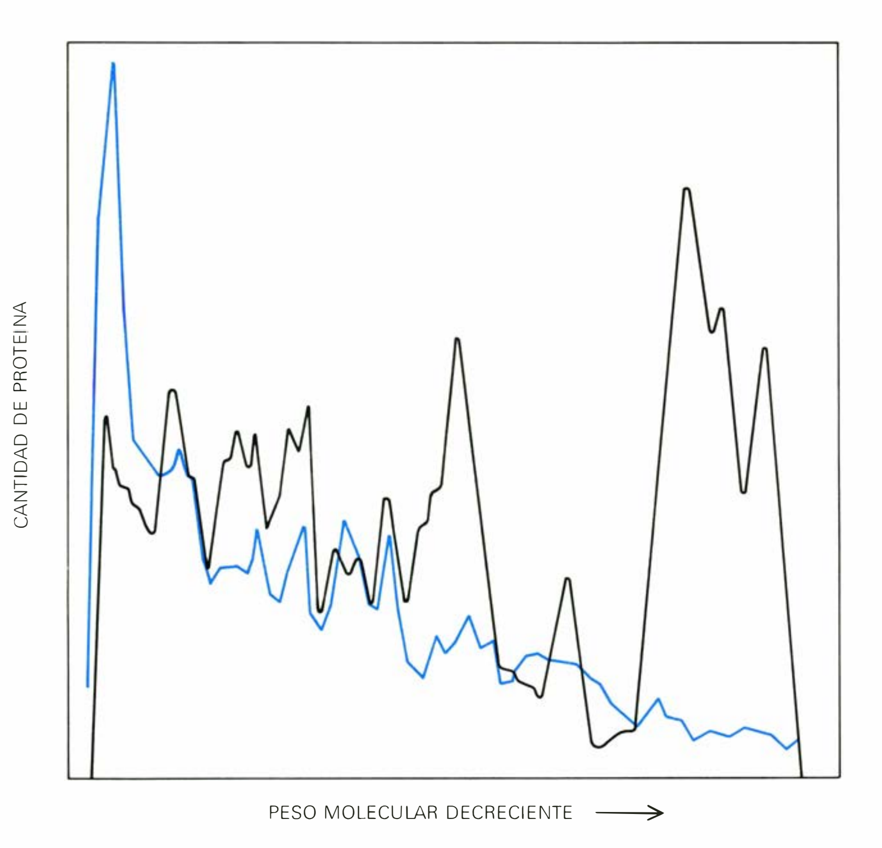




HAY UN ESPECTRO DISTINTO para cada tejido diferente de proteínas no histónicas. Cada curva registra las fluctuaciones en la densidad óptica y, por lo tanto, en la cantidad de proteína, a lo largo de un gel como el de la derecha de la página opuesta. El modelo de picos, y por consiguiente el de cada proteína no histónica, varía en los geles que contienen dicho tipo de proteínas, que han sido extraídas de cuatro cortes histológicos especializados y diferentes de un múrido.

nifiestan falta de control del crecimiento, lo cual es típico de las formas malignas. Los últimos resultados de nuestros laboratorios indican que hay cambios en la composición y metabolismo de las proteínas no histónicas, cromosómicas, casi inmediatamente después de la infección de las células con virus tumorales. Estas modificaciones pasan a ser permanentes en las células que se transforman y permanecen enfermas. Hasta ahora, no hay una prueba definitiva de una relación de causa a efecto, pero parece razonable postular que las proteínas no histónicas tengan una función importante en la mediación de las modificaciones inducidas por los virus en la expresión génica y que están asociadas con la condición neoplásica.





LA VELOCIDAD DE SINTESIS de ciertas proteínas no histónicas, y por tanto su cantidad respectiva, varía en un mismo tipo de célula en distintos momentos de diferenciación. Las curvas en color registran las velocidades de síntesis y, las curvas negras, las cantidades de las proteínas no histónicas localizadas en las células diferenciadas de un neuroblastoma de ratón *(véase el gráfico superior)* **y en las no diferenciadas** *(véase el gráfico inferior).* **En el estado diferenciado se manifiesta una reducción selectiva en la velocidad de síntesis de las proteínas de elevado peso molecular.**

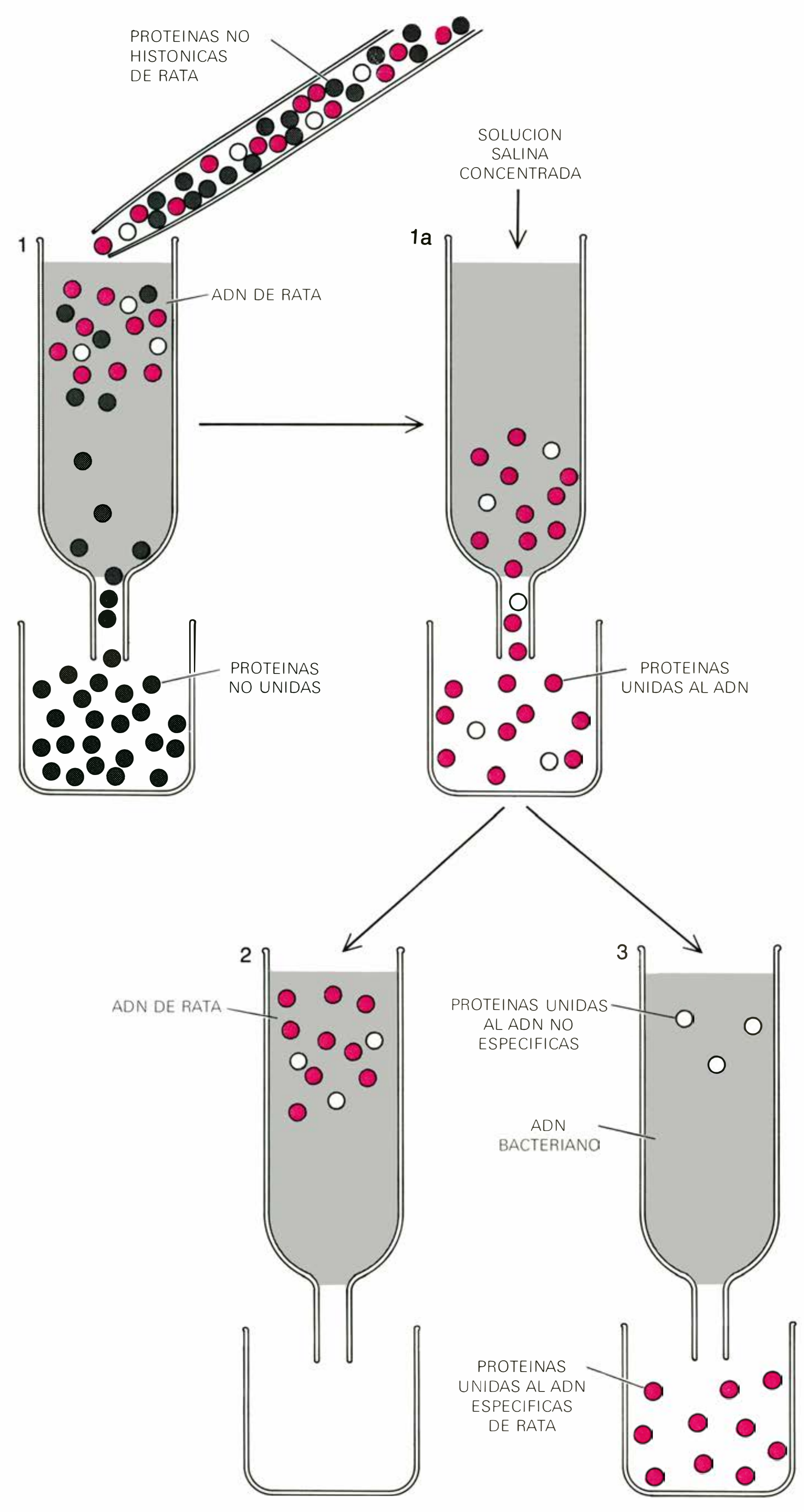
Otra propiedad peculiar de las proteínas cromosómicas no histónicas, que apoya la idea de que actúan en la regulación génica específica, es la de su capacidad de ligarse al ADN. En las células bacterianas, en las que se han identificado y aislado las proteínas reguladoras específicas, se ha puesto en claro que funcionan enlazándose a secuencias específicas en regiones adecuadas del ADN. De manera análoga podría esperarse que las moléculas reguladoras específicas, presentes en las células eucariotas, se unan a tipos particulares de ADN. En 1970, los experimentos llevados a cabo en uno de nuestros laboratorios (el de Kleinsmith) demostraron, por primera vez, que algunas proteínas no histónicas se unen específicamente al ácido desoxirribonucleico.

Tales experimentos utilizaron la técnica de cromatografía sobre celulosa del ADN desarrollada por Bruce M. Alberts, de la Universidad de Princeton. En dicha técnica se absorbe el ADN en una matriz de celulosa que se encierra en una columna de vidrio, por la cual se hacen pasar las proteínas objeto de análisis. Las proteínas que se unan al ADN quedan atrapadas en la columna; pueden extraerse luego para su examen. Si, por una de estas columnas que contenga ADN de rata, se hacen pasar proteínas no histónicas, preparadas a partir de hígado de rata, se fija un pequeño porcentaje de proteínas. Al extraer de la columna las proteínas unidas al ADN, y hacerlas pasar por una nueva columna con ADN bacteriano, la mayoría de ellas la atraviesan libremente *(véase la ilustración de la página 39)*. En otras palabras, algunas proteínas no histónicas son capaces de reconocer secuencias específicas de nucleótidos del ADN, de modo que se unen a su propio ADN pero no a uno extraño. Una tal interacción específica entre estas proteínas y el ADN es exactamente lo que cabría esperar si algunas se hallaran comprometidas en la regulación de la actividad de genes específicos. Esta unión específica del ADN y las proteínas no histónicas quedó confirmada por las observaciones registradas en otros laboratorios, de los cuales recordamos los de Allfrey, Bonner y Lubomir S. Hnilica de la Universidad de Texas en Houston.

Pruebas experimentales

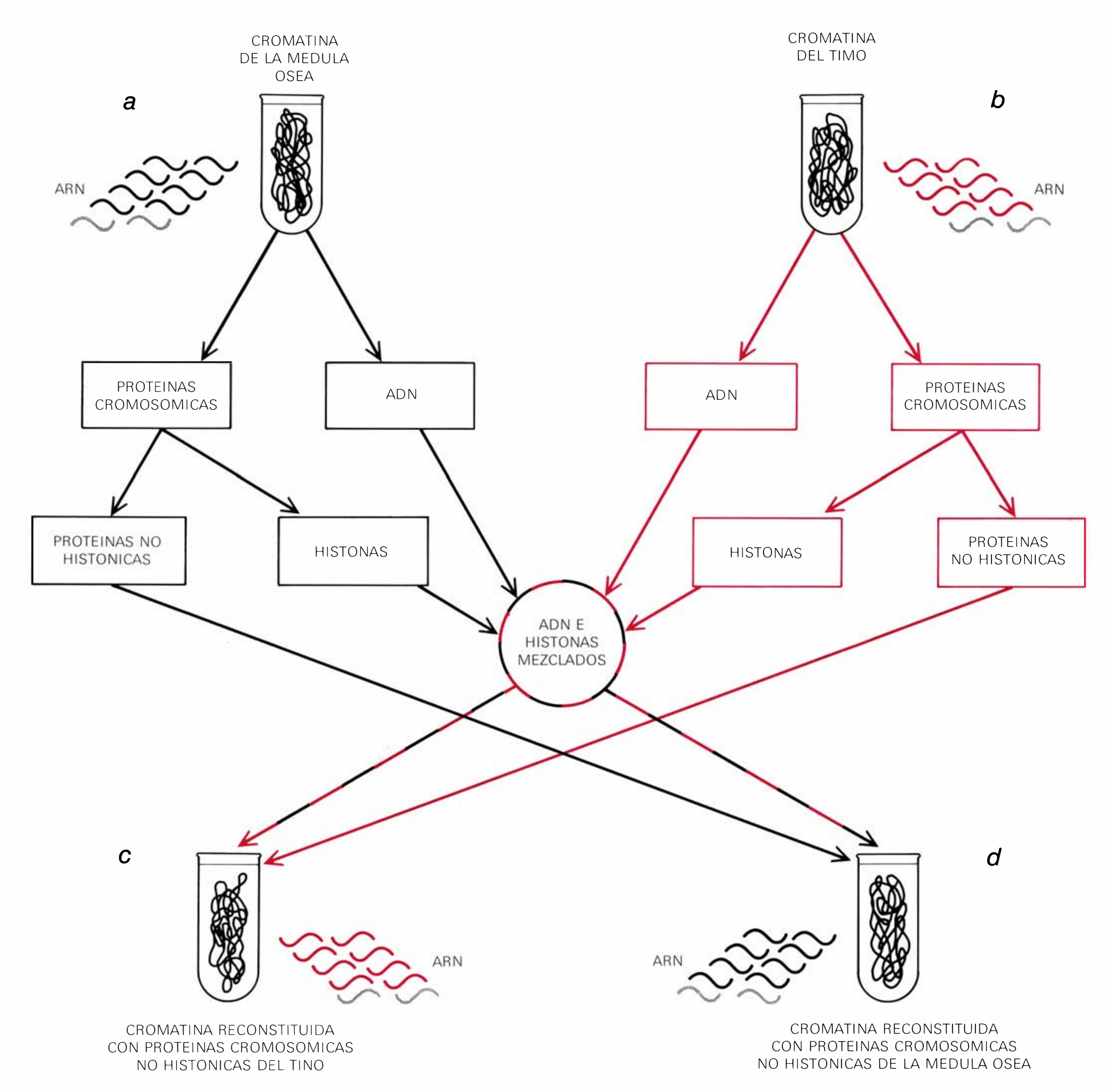
Para determinar directamente si la presencia de proteínas no histónicas puede alterar la transcripción génica, es necesario utilizar sistemas en los que puedan medirse con exactitud los efectos de las proteínas no histónicas sobre la síntesis de ARN. En 1968, un informe presentado por el laboratorio de Tung-Yue Wang, de la Universidad estatal de Nueva York, en Buffalo, indicaba que la adición de proteínas no histónicas a un sistema libre de células que contuviese ADN purificado y ARN polimerasa podía hacer revertible la inhibición de la síntesis de ARN, producida normalmente por la presencia de histonas. Estos resultados eran coherentes con un primer informe de Thomas A. Langan, que estaba trabajando entonces en la Rockefeller University; según dicho informe, cuando la histona forma un complejo con la proteína no histónica, aquélla sólo ejerce un efecto inhibitorio parcial en la síntesis de ARN. Ching-Sung Teng y Terrell H. Hamilton, de la Universidad de Texas, en Austin, y Thomas C. Spelsberg y Hnilica, de Houston, emplearon aproximaciones similares. Ambos grupos encontraron, en diferentes sistemas, que las proteínas cromosómicas no histónicas impedían la inhibición completa de la síntesis de ARN por parte de las histonas. El grupo de Wang fue más allá, demostrando que el aumento de la transcripción producido por la adición de proteínas no histónicas representaba la activación de regiones del genoma que se habían reprimido antes. En conjunto, estos estudios demostraron claramente que las proteínas cromosómicas no histónicas pueden influir en la transcripción. De todos modos, la interpretación de los resultados se hizo difícil, dado que en un sistema libre de células los componentes de la cromatina estaban mezclados en condiciones que quizá no les permitan asociarse como lo están en el núcleo de una célula intacta.

La prueba más directa de que las proteínas cromosómicas no histónicas desarrollan una función en la regulación de la transcripción tisular específica se desprende de los experimentos de reconstitución de la cromatina, en los que se separan los distintos componentes de la cromatina y se vuelven a juntar en diferentes combinaciones. En los primeros experimentos llevados a cabo siguiendo esta orientación, R. Stewart Gilmour y John Paul, del Beatson Institute for Cancer Research de Glasgow, aislaron las cromatinas del timo y de la



LA UNION ESPECIFICA entre proteínas no histónicas y ADN se demostró haciendo pasar proteínas no histónicas de hígado de rata a través de una columna de celulosa en la que se había introducido, previamente, ADN de rata. Unas proteínas se detenían en la columna que había absorbido ADN; las proteínas no unidas *(en gris)* **salían fuera** *(1).* **Las proteínas unidas** *(en color y en blanco)* **se extrajeron con una solución salina concentrada** *(1a).* **A éstas se les hizo pasar entonces a través de columnas que contenían ADN de rata o ADN de la bacteria** *Escherichia coli.* **Todas las proteínas se unieron al ADN de rata** *(2);* **sólo una pequeña proporción lo hizo con el ADN de** *E. coli (3).* **Las que se unieron sólo al ADN de rata pero no al de** *E. coli* **eran específicas del ADN de rata.**

médula ósea de conejo y las disociaron hasta ofrecer por separado el ADN, las histonas y los componentes no histónicos. Reconstituyeron después las cromatinas sirviéndose de un método desarrollado por Huang y Bonner. En el experimento de control se volvió a conjuntar simplemente el ADN, las histonas y las proteínas no histónicas del timo de conejo. Entonces se dispuso la cromatina reconstituida de forma que sirviese de molde para la síntesis de ARN en presencia de ARN polimerasa. En las pruebas verificadas para calibrar su capacidad de unión al ADN extraído del timo o de la médula ósea de conejo, el ARN transcrito se comportó exactamente igual que el ARN de timo sintetizado normalmente. Gilmour y Paul mezclaron luego el ADN y las histonas de timo y de médula y reconstituyeron la cromatina añadiendo a la mezcla proteínas no histónicas del timo. El ARN fabricado por esta cromatina híbrida se comportó otra vez como el ARN del timo, aunque la cromatina contenía ADN e histonas de la médula ósea. Por otra parte, cuando se reconstituyó el ADN y las histonas mezclados con proteínas no histónicas de la médula ósea, el ARN

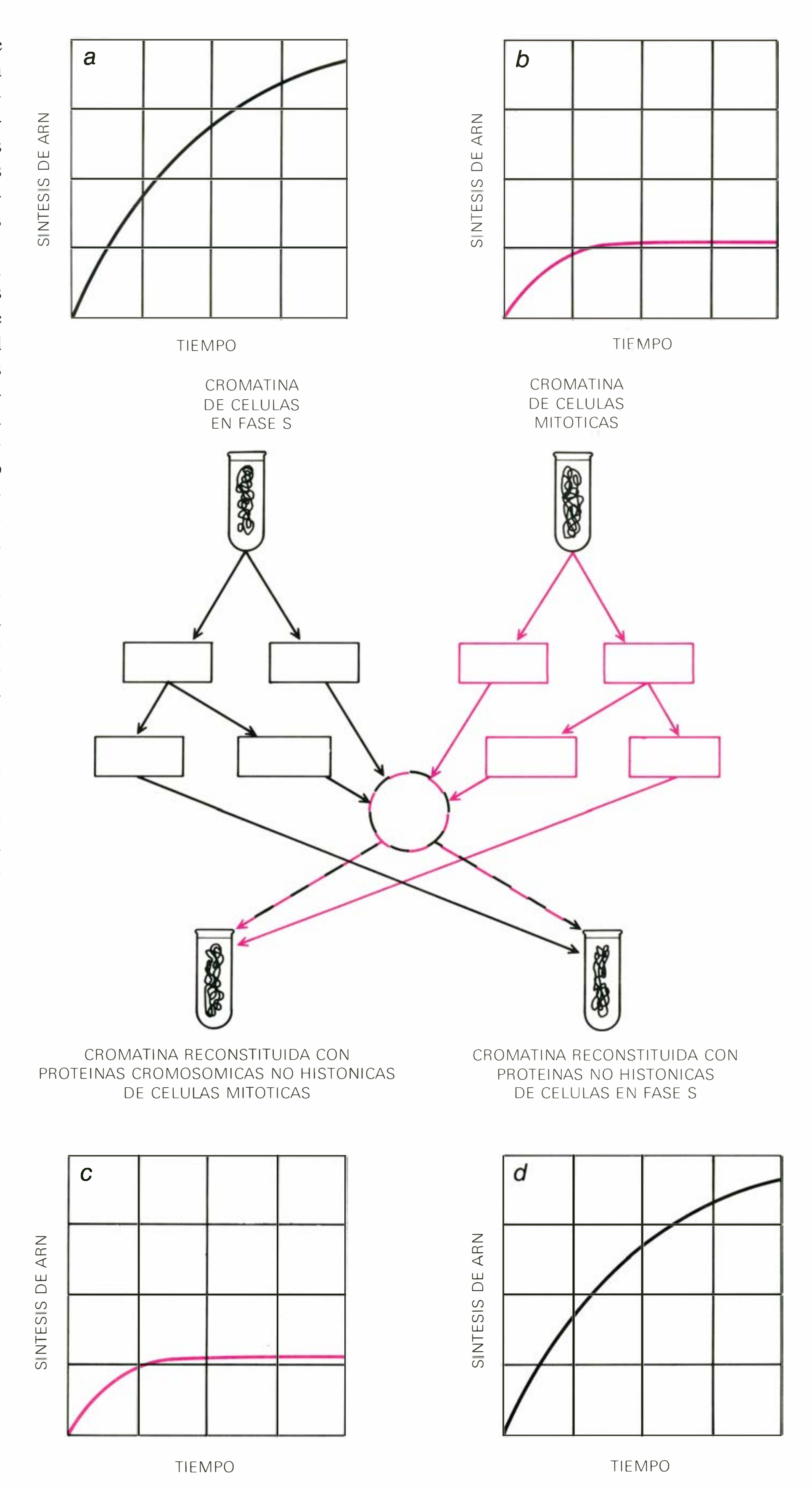


EL EXPERIMENTO DE RECONSTITUCION DE CROMATINA llevado a cabo por R. Stewart Gilmour y John Paul, del Beatson Institute for Cancer Research, puso en evidencia que las proteínas no histónicas eran responsables de la transcripción específica del tejido. Se aisló la cromatina de células de timo y médula ósea de conejo. Cada cromatina formó un ARN característico *(a, b)*. Cuando cada cromatina se disoció en ADN, histonas y fracciones no histónicas, y se volvió a constituir, formó el mismo ARN. Luego se disociaron las cromatinas, y se mezcló el ADN con las histonas procedentes de ambas preparaciones. Al combinarse la mezcla de ADN e histonas con las proteínas no histónicas del timo, la cromatina reconstituida produjo un ARN que se comportó como el ARN del timo *(c)*. Al combinarse la mezcla de ADN e histonas con las proteínas no histónicas de la médula, la cromatina reconstituida produjo un ARN que se comportó como el de la médula *(d)*.

formado por la cromatina híbrida se comportó como el ARN de la médula ósea *(véase la ilustración de la página anterior).* Los experimentos pusieron en claro que la presencia de las proteínas no histónicas, específicas de los tejidos, determina exactamente qué genes se transcribirán en los diferentes tejidos.

¿Qué ocurre en un mismo tejido en diferentes momentos? ¿Son responsables también las proteínas no histónicas de los cambios transitorios en la actividad transcriptiva, asociados a los cambios de estado metabólico, en un tipo de célula dado? El ciclo celular proporciona un modelo ideal para investigar directamente la influencia de las proteínas no histónicas sobre la actividad transcriptiva del genoma, puesto que la cromatina aislada a partir de células en fase *S* (período de replicación activa del ADN) tiene mayor actividad en la síntesis de ARN que la cromatina aislada a partir de células en mitosis. Los experimentos de reconstitución de la cromatina, llevados a cabo por uno de nosotros (Gary Stein) y por John Farber, en la Temple University, demostraron que las proteínas no histónicas son las responsables de esta modulación de la transcripción génica durante el ciclo celular. La cromatina, reconstituida a partir de la mezcla de ADN e histonas con proteínas no histónicas que procedían de células en mitosis, mostraba menos capacidad para la síntesis de ARN que la cromatina reconstituida en presencia de proteínas no histónicas que procedían de células en fase *S (véase la ilustración de la derecha).* Para eliminar la posibilidad de que estuviesen implicadas las histonas en estas alteraciones de la transcripción génica, también se llevaron a cabo experimentos de reconstitución en presencia de histonas procedentes de células en fase *S* o de células en mitosis. En este último supuesto, las actividades transcriptivas de los dos tipos de cromatinas reconstituidas fueron idénticas, demostrando que las histonas no establecen variaciones en la capacidad de las secuencias de ADN para la síntesis de ARN durante el ciclo celular.

Para demostrar con claridad que las proteínas no histónicas regulan la actividad de genes determinados, es preciso someter a prueba la transcripción de un gen específico. En los últimos años ha sido posible sintetizar en el laboratorio copias radiactivas de genes aislados, "retrocediendo" desde un ARN conocido hasta su ADN complementario Estas copias radiactivas de genes se



OTRO EXPERIMENTO DE RECONSTITUCION demostró que las proteínas no histónicas dirigen las variaciones de transcripción que se advierten en los distintos estadios del ciclo celular. La cromatina extraída de células en fase *S* sintetiza ARN a mayor velocidad *(d)* que la cromatina extraída de células en mitosis *(b).* Se disociaron ambas cromatinas, y se mezcló el ADN con las histonas, como en el experimento ilustrado en la página opuesta. La cromatina reconstituida con proteínas no histónicas procedentes de células en mitosis sintetizó ARN al ritmo propio de la cromatina de las células mitóticas *(c).* La cromatina reconstituida con proteínas no histónicas procedentes de células en fase *S* produjo ARN a un ritmo superior, propio de la mencionada fase *S*.

ARN MENSAJERO

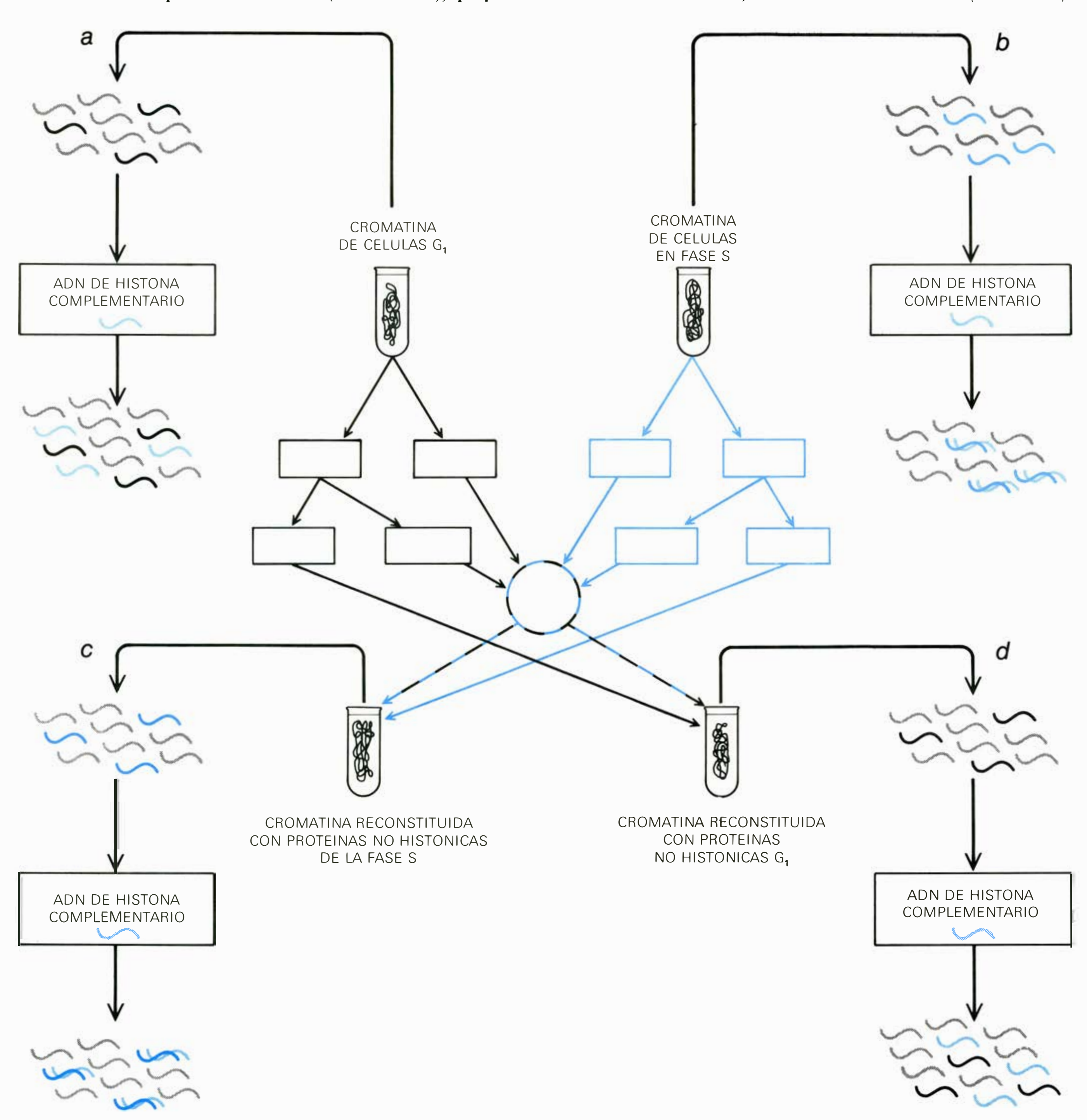
HIBRIDO

ADN COMPLEMENTARIO

ADN POLIMERASA
DEPENDIENTE DEL ARN

HIDROLISIS

LOS EXPERIMENTOS DE HIBRIDACION sirven para identificar productos específicos del gen. Se aísla un ARN conocido *(en color oscuro).* **En presencia del enzima ADN polimerasa dependiente del ARN ("transcriptasa inversa"), el ARN sirve de molde para la formación de una cadena complementaria de ADN** *(en color claro),* **que puede marcarse por la incorporación de subunidades radiactivas. El híbrido ARN - ADN se disocia por hidrólisis, y se libera un ADN radiactivo complementario, un gen sintetizado. Este puede servir como prueba para identificar el ARN que se forma de él; en efecto, cuando se añade a una solución de ARN, hibridará sólo con este ARN** *(a la derecha).*



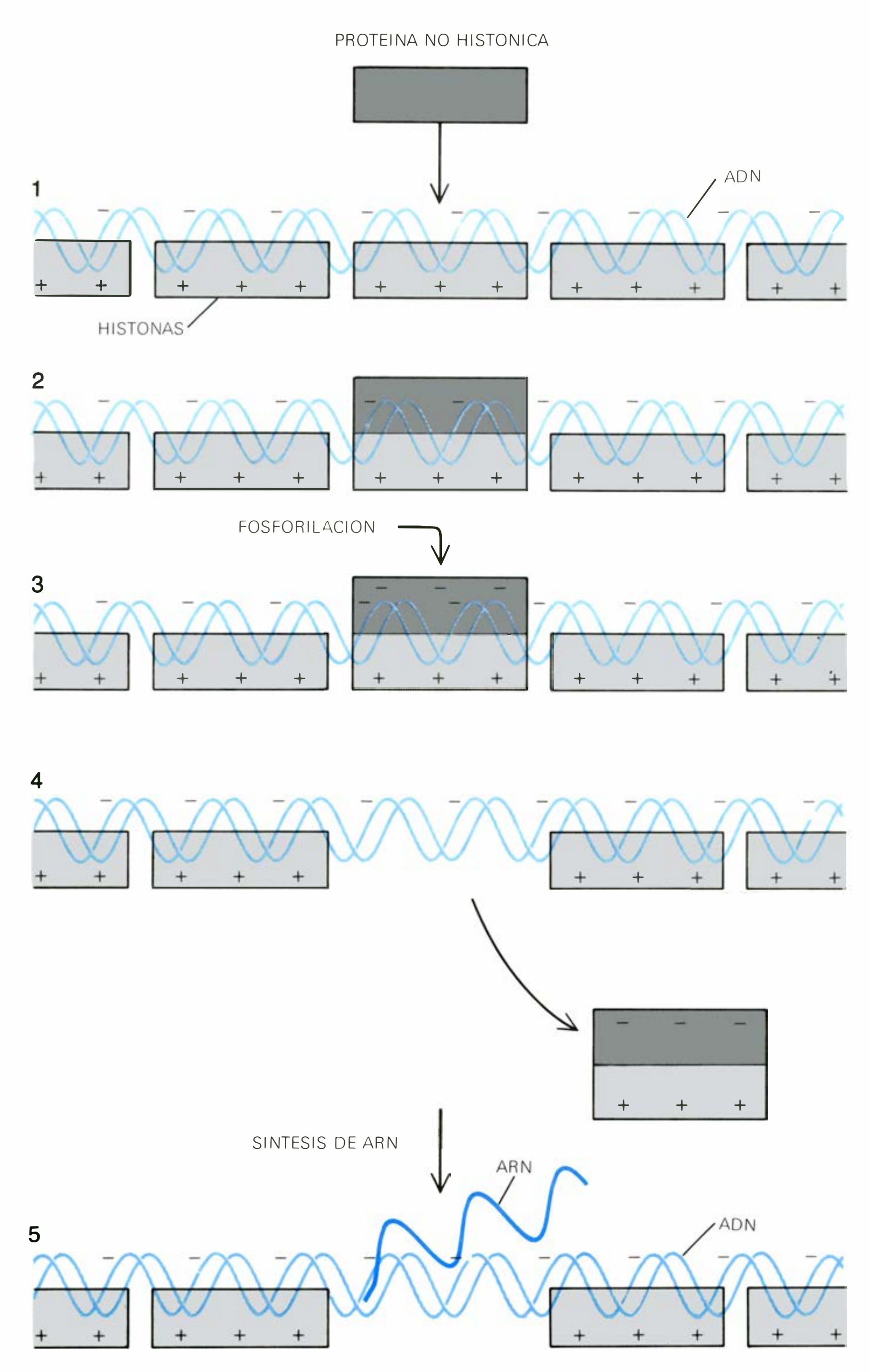
RECONSTITUCION E HIBRIDACION combinadas en un mismo experimento. Se preparó un ADN radiactivo complementario al ARN de las histonas, y se usó después como prueba *(en color claro)* **para demostrar que, mientras que la cromatina de las células en fase G_1 no forma ARN de las histonas** *(a),* **sí lo hace la cromatina de células en fase** *S (en color oscuro, b).* **Se disociaron las cromatinas, se mezclaron el ADN y las histonas y se reconstituyeron después con los dos tipos de proteínas no histónicas. Se prepararon ARN a partir de las dos cromatinas reconstituidas, y se probaron con el ADN. La cromatina, reconstituida con proteínas no histónicas de la fase** *S,* **sintetizó ARN de histonas tal como lo hizo la cromatina de la fase** *S* **natural** *(c).* **La cromatina, reconstituida con proteínas no histónicas en fase G_1** *(d),* **no lo hizo.**

unirán específicamente a sus ARN complementarios para los cuales codifican y de este modo pueden usarse como "sondas" sensibles para evidenciar la presencia de ARN que se ha transcrito a partir de genes particulares *(véase la ilustración de la cabecera de la página 42).*

Gilmour y Paul emplearon recientemente esas "sondas" complejas en una serie apasionante de experimentos sobre reconstitución de la cromatina. Esta se aisló a partir de hígado de feto de ratón, que sintetiza normalmente globina (la proteína de la hemoglobina), y de cerebro de ratón, que no sintetiza globina. Se transcribió cada preparación de cromatina en un sistema libre de células, caracterizándose al ARN sintetizado mediante la comprobación de su capacidad para unirse a copias radiactivas de los genes que codifican para la globina. Conservando el mismo tipo de transcripción observado en células intactas, la cromatina de hígado fetal sintetizó ARN que codificaba para la globina mientras que la cromatina de cerebro no. Gilmour y Paul prosiguieron luego sus investigaciones disociando las cromatinas y reconstituyéndolas en presencia de distintos tipos de proteínas no histónicas. Vieron que, cuando se reconstituía la cromatina de cerebro en presencia de proteínas no histónicas obtenidas de tejido cerebral, ella no sintetizaba ARN para la globina, mientras que si la reconstitución se llevaba a cabo en presencia de proteínas no histónicas de hígado fetal, se transcribía el gen para la globina. Tales resultados corroboran firmemente la idea de que la presencia de proteínas no histónicas específicas determina que se transcriba el gen de la globina en el tejido hepático fetal y no suceda lo mismo en el tejido cerebral.

Con una metodología y un enfoque muy parecidos, nuestro grupo de Florida demostró recientemente que las proteínas cromosómicas no histónicas son las responsables del retraso en la transcripción de los genes de las histonas hasta la fase *S* del ciclo celular *(véase la ilustración inferior de la página opuesta.*

Mediante una "sonda" histona-ADN radiactivo destinada al examen del ARN formado en el tubo de ensayo, se demostró primero que la cromatina procedente de células en fase *S* transcribía ARN para las histonas, mientras que la cromatina obtenida de células en fase G_1 del ciclo celular (período anterior a la replicación del ADN) no lo hacía. Este resultado concordaba con una comprobación precedente de que la síntesis de histonas y la replicación del ADN están acopladas. Llegamos, más tarde, a demostrar que la cromatina reconstituida con proteínas no histónicas procedentes de células en fase *S* transcribía los ARN para las histonas en la misma proporción que la cromatina natural procedente de células en fase *S*. Contrariamente, la cromatina reconstituida con proteínas no histónicas aisladas de la cromatina de células en fase G_1 no mostraba ninguna transcripción identificable de ARN de las histonas. Todo lo cual viene a significar que las proteínas cromosómicas no histónicas dirigen la elaboración de información procedente de genes específicos durante períodos bien definidos del ciclo de la célula.



SE ACTIVA UN GEN, de acuerdo con un modelo hipotético, cuando una proteína no histónica reconoce *(1)* un sitio *(locus)* específico en un segmento del ADN que está generalmente reprimido por las histonas. La proteína no histónica se une al sitio *(2)*, se fosforila, y, por tanto, se carga negativamente *(3)*. En consecuencia, repele al ADN cargado negativamente y se va con la histona cargada positivamente, dejando desnudo el sitio *(locus)* del ADN *(4)*. Dicho sitio, que ya no está reprimido por la histona, queda disponible para transcribir ARN *(en color oscuro, 5)*.

Un mecanismo posible

Si se acepta que las proteínas cromosómicas no histónicas desempeñan una función clave en la regulación de la transcripción de genes individuales, en distintas células especializadas y en períodos diferentes en un mismo tipo celular, podemos preguntarnos cómo se verifica tal control. ¿Cuál es el mecanismo molecular fundante? Una propiedad notable de las proteínas no histónicas que nos sugiere un camino de estudio es su vasta fosforilización. En una serie de experimentos comenzados en 1966, Kleinsmith (firmante de este artículo), Allfrey y Mirsky observaron que las proteínas no histónicas se están modificando continuamente en el núcleo por la adición y eliminación de grupos fosfato. Pensamos que, dado que los grupos fosfato están cargados negativamente, semejantes modificaciones podrían alterar la estructura y las interacciones funcionales de las proteínas no histónicas en la cromatina. Desde estas primeras observaciones, los investigadores de muchos laboratorios, estudiando diferentes sistemas biológicos, han notado que la fosforilización de las proteínas no histónicas va asociada a la diferenciación celular y a la activación de los genes. Además, se ha demostrado que la capacidad de las proteínas no histónicas para estimular la síntesis de ARN en sistemas libres de células depende de su estado de fosforilización.

Tales observaciones sugieren que la fosforilización de las proteínas no histónicas puede estar implicada de algún modo en el mecanismo gracias al cual estas moléculas regulan la transcripción génica. Un modelo propuesto postula la unión con un ADN específico y la fosforilización *(véase la ilustración de la página 43)*. Creemos que una proteína no histónica específica se une a un sitio *(locus)* específico del ADN, reprimido por las histonas. En ese punto, se fosforila la proteína no histónica, y dado que los grupos fosfato se hallan dotados de carga negativa, dicho polipéptido comienza a repeler al ADN, cargado asimismo negativamente, al mismo tiempo que se asocia intensamente con las histonas, cuya carga es positiva. Como resultado de estos dos procesos, se origina la liberación, respecto al ADN, del complejo histona-proteína no histónica. El área desnuda del ADN resultante, libre ya de su histona inhibidora, puede transcribirse para dar ARN.

Aunque este modelo de liberación o despegue respecto de la histona se desenvuelve ampliamente en el terreno de la hipótesis, ciertas observaciones se muestran coherentes con sus líneas generales. Según quedó dicho, los genes se transcriben más intensamente durante la fase *S* que durante la mitosis. Recientemente hemos podido observar que las proteínas no histónicas se fosforilizan también más intensamente en la fase *S* que durante la mitosis y, asimismo que las histonas se unen con menor fuerza al ADN durante la fase *S* que durante la mitosis. Nuevos experimentos sobre reconstitución de cromatina, realizados en nuestro laboratorio de Florida, han demostrado que las proteínas no histónicas asociadas con el genoma en la fase *S* son responsables específicamente de la unión menos enérgica de las histonas al ADN en ese período del ciclo celular. Resultados todos ellos que vienen a indicar, si no otra cosa, por lo menos una estrecha correlación entre la actividad génica, la fosforilización de las proteínas no histónicas y la unión de las histonas al ADN.

El modelo de liberación de las histonas es sólo uno de los muchos propuestos para justificar el papel de las proteínas cromosómicas en la regulación génica. Aunque no se han comprobado los pormenores del mecanismo, la comprensión de las proteínas cromosómicas y sus interacciones con el ADN va progresando por días. No cabe duda de que las histonas y las proteínas no histónicas desarrollan importantes misiones en la determinación de las propiedades estructurales y funcionales del genoma. Mientras que las histonas están implicadas en el mantenimiento de la estructura de la cromatina y en la represión no específica de secuencias genéticas, parece evidente que las proteínas cromosómicas no histónicas reconocen sitios *(loci)* genéticos definidos y, mediante ello, regulan la transcripción de información genética específica.

Todo hace presumir que en breve se aislarán las proteínas que regulan la expresión de genes específicos, y se abrirá la puerta a la posibilidad de un cierto tipo de ingeniería genética: podrán inocularse proteínas en las células para modificar las anormalidades en la transcripción génica asociadas con el desarrollo, la diferenciación y una amplia gama de enfermedades, entre ellas el cáncer. Dicha manipulación revolucionaría las posibilidades del hombre ante determinados trastornos hondamente degenerativos. Pero, como en cualquier otro caso de bioingeniería humana, tal enfoque plantea graves y complejos problemas éticos que requieren una consideración muy detenida.

Ciencia y sociedad

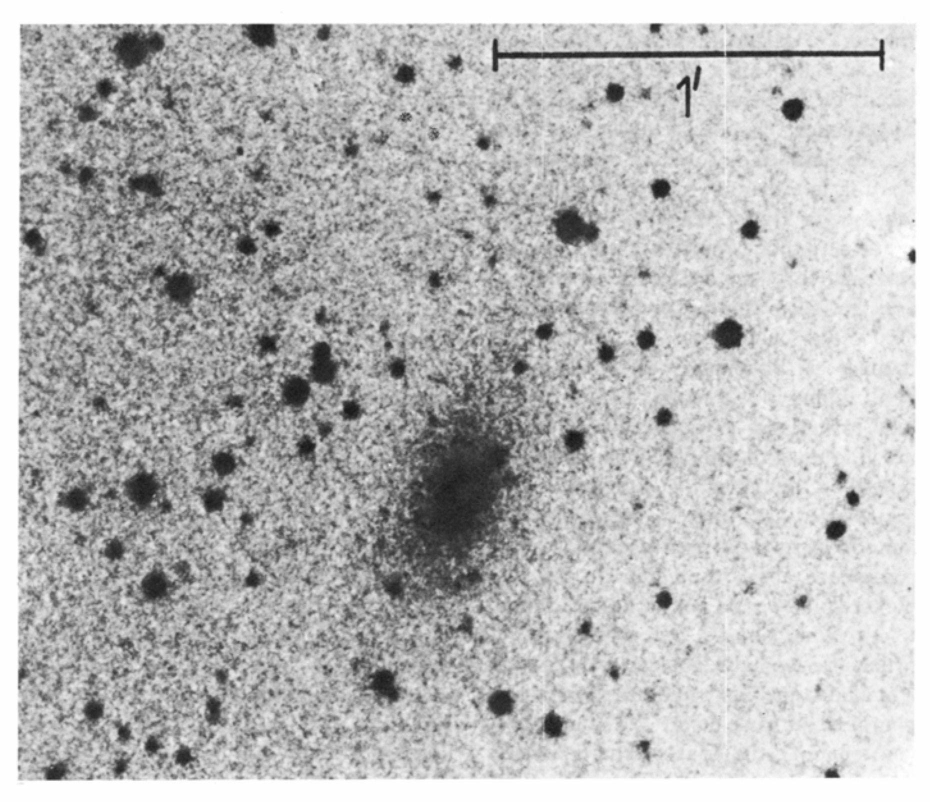
Descubrimiento de la galaxia cercana Calar Alto I

El Dr. Antonio Romañá, ex director del Observatorio del Ebro (Tortosa) nos informa sobre ese descubrimiento del Observatorio de Calar Alto y su incidencia en la astronomía.

El 28 de mayo del año en curso, en la conferencia de prensa celebrada en Almería por el Consejo de Administración de Observatorio Hispano-Alemán de Calar Alto, anunció el Prof. H. Elsässer, Director del Max-Planck Institut für Astronomie, de Heidelberg, el descubrimiento desde el nuevo observatorio de una nueva galaxia, a la que se había designado con el nombre de Calar Alto I. El anuncio al mundo científico se había hecho unos meses antes. El nuevo descubrimiento de una galaxia no es noticia sensacional. Existen en número de millones, centenares de millones incluso; y, a medida que se va disponiendo de telescopios más potentes, ópticos o radioeléctricos, crece sin cesar el número de las descubiertas en las profundidades del espacio. La trascendencia del descubrimiento de la Calar Alto I está, por una parte, en la muy probable proximidad de la misma, y, por otra, en la técnica utilizada para su hallazgo. Lo cual puede dar lugar a nuevos descubrimientos que modifiquen, tal vez, nuestros conocimientos sobre el universo.

Las galaxias se distribuyen formando grupos o enjambres. La Vía Láctea, a la que nosotros pertenecemos, está incluida en uno de estos grupos, el llamado Grupo Local, constituido por unas 27 galaxias cuya distancia es inferior a cinco millones de años de luz. Este grupo se creía enteramente conocido y compuesto por dos galaxias espirales gigantes, la nuestra y la de Andrómeda, dos espirales de tipo medio, la del Triángulo y la Gran Nube de Magallanes, ocho irregulares enanas, once elípticas enanas o "dispersas", cuatro objetos intermedios entre galaxias y enjambres globulares, y tal vez dos enjambres globulares más, perdidos en el espacio. Es un grupo bastante disperso sin condensación central, compuesto quizás por dos subgrupos, centrado el uno sobre la nebulosa de Andrómeda y el otro sobre la nuestra, y ocupando en total un volumen de unos cien trillones de años de luz cúbicos. Su masa total se cree ser 650 mil millones de veces la del Sol. El descubrimiento de una nueva galaxia próxima puede no obstante cambiar el panorama, máxime si este descubrimiento tiene visos de ser preludio de otros.

En efecto, la posición excéntrica de nuestro Sistema Solar dentro de la Vía Láctea y la gran cantidad de cuerpos celestes y masas de polvo cósmico que se concentran cerca del plano ecuatorial de nuestra galaxia, hace que sea muy difícil la visión de lo que hay al otro lado de la misma y que fuese hasta hace poco prácticamente imposible la fotografía de tales objetos. De aquí que, siendo fácil el descubrimiento de nuevas galaxias en las direcciones alejadas del plano ecuatorial de nuestra galaxia, fuese dificilísimo en las inmediaciones del mismo. Hace algunos años se comenzó a desarrollar la llamada astronomía del infrarrojo, gracias a la cual se pudieron fotografiar cuerpos celestes antes inabordables por el ojo humano; tal, por ejemplo, la galaxia oculta Maffei I, perteneciente al grupo Osa Mayor-La Jirafa, situado a unos ocho millones de años de luz. Al decidir la Max-Planck Gesellschaft la creación del Observatorio de Calar Alto, concibió el empleo de nuevas técnicas de infrarrojo y la construcción de aparatos apropiados para ello, gracias a los cuales se ha conseguido fotografiar la Calar Alto I. Mediante el empleo de una cámara de tubo de imagen con cátodo Sl, se fotografiaron un par de objetos difusos dentro del área S119, una extensa región de emisión en el Cisne, cerca de la nebulosa Norteamérica, tras una exposición de 35 minutos. Los resultados se deducen a partir de fotografías en IR en longitud de onda $\lambda = 0{,}92\ \mu m$, hechas con el telescopio de 1,23 m. En la foto inferior se aprecia una condensación nebulosa elíptica irregular inclinada del ángulo superior derecho hacia el inferior izquierdo, de más de medio minuto de arco, rodeada de un débil halo que está afectado por la emisión de la región S119. Debido a su núcleo distinto y a este halo, la nebulosa tiene un cierto aspecto de galaxia lejana. Pero su comparación con la galaxia Maffei I, y, en especial, las medidas densitométricas a lo largo de sus ejes mayores hacen muy probable que se trate de una galaxia cercana. El segundo



La nebulosa Calar Alto I descubierta en el nuevo Observatorio Hispano-Alemán de Calar Alto.

objeto es de naturaleza mucho más incierta y puede ser un enjambre globular galáctico obscurecido u otro objeto galáctico, sin que se excluya que se trate también de una galaxia. Varios grupos astronómicos se han interesado mucho en ello y en el Observatorio de Lick la están estudiando radioastronómicamente y confían poder determinar pronto su velocidad radial y su distancia. Por una porción de indicios se puede llegar a la conclusión de que no es una galaxia joven. Hace unos meses, Christian Simonson II, de la Universidad de Maryland, descubrió lo que parece ser una galaxia enana, más próxima que las Nubes de Magallanes y satélite de nuestra Vía Láctea, estudiando la distribución en el cielo del hidrógeno, deducida de las observaciones radioeléctricas en longitud de onda de 21 cm. Notó la existencia de una cadena peculiar de nubes de hidrógeno con grandes velocidades radiales de aproximación que se extendía desde Orión, a través de Géminis y Auriga, casi paralelamente al plano central de la Vía Láctea, ligeramente hacia el norte del mismo. Simonson las interpretó como componentes de una galaxia próxima, ópticamente invisible por estar enmascarada por las nubes de polvo y densos campos estelares de la Vía Láctea. Por su gran tamaño angular dedujo Simonson que se tenía que tratar de un cuerpo celeste muy cercano; por comparación de la forma observada con la deformación que debería producir en ella la atracción de la Vía Láctea a diversas distancias, le asignó Simonson una distancia de 55.000 años de luz, menos de un tercio de la distancia de las Nubes de Magallanes; en cuanto a su masa, por las velocidades de rotación de su cuerpo central, la creyó del orden de 1/200 de la masa de la Vía Láctea. Su estudio se presenta muy difícil, por tener que aparecer forzosamente sus estrellas, cuando se consiga fotografiarlas, mezcladas con las de la Vía Láctea, casi cien veces más numerosas. Simonson y Grayzeck, de la Universidad de Nevada, en las Vegas, han formulado y emprendido un plan de ataque. Considerando la forma en que se ha descubierto la Calar Alto I, es lógico deducir que la aplicación del método empleado podría ayudar poderosamente a la resolución de los problemas planteados por el descubrimiento de Simonson, y, en todo caso, es prueba de los hallazgos que cabe esperar a medida que las regiones próximas al plano ecuatorial de la Vía Láctea vayan siendo estudiadas sistemáticamente desde el nuevo Observatorio Hispano-Alemán.

La trascendencia de tales hallazgos puede ser muy grande. De todos es bien conocido el papel que el conocimiento de la densidad media del Universo juega en las diferentes teorías cosmológicas; y para el cálculo de ρ_0 un buen punto de partida es el conocimiento de la den sidad media del Grupo Local. Si descubrimientos como el de Calar Alto I se multiplicaran, resulta evidente que este valor inicial sería otro y que, por tanto, habría que introducir probablemente algunos retoques y modificaciones en las teorías cosmológicas actualmente más en boga.

Un nuevo eslabón

El nuevo cráneo humanoide descubierto por Amini Mturi en las excavaciones del lago Ndutu, al norte de Tanzania, presenta unas peculiaridades muy interesantes que lo constituyen en presumible eslabón entre el *Homo erectus pekinensis* y el *Homo sapiens,* ya que posee rasgos de uno y otro.

No puede clasificarse como *Homo sapiens;* tampoco, a pesar de su estrecho parecido, como *Homo e. pekinensis.* Es más, por manifestar características más avanzadas que este último y hallarse ubicado geográficamente en Africa y no en Asia, permite la creación de una nueva subespecie de *Homo erectus,* a la que él perteneciera.

R. J. Clark ha publicado en *Nature* los rasgos comparativos que asimilan el cráneo descubierto a una y a otra especie de *Homo.* Con el *H. e. pekinensis* coinciden en lo siguiente: la forma y la curva del occipital, cuya prominencia nucal es muy destacada; la forma de la región mastoidea; la verticalidad frontal; la prominencia supraorbital insinuada; el gran espesor de la bóveda; y el perfil, visto en norma vertical. Difiere del *H. e. pekinensis* y se aproxima al *Homo sapiens* en estos otros caracteres: la presencia en el adulto de protuberancias parietales muy marcadas; los flancos de la bóveda más verticales, visto en norma occipital; la presencia de un proceso estiloideo osificado; la ausencia aparente de una prominencia sagital; y la no extensión de la cresta supramastoidea por encima del meato auditivo externo.

El cráneo, que se conserva en el Museo Nacional de Tanzania, en Dar es Salaam, mantiene casi completos el occipital y el temporal izquierdo; ha perdido la fosa glenoidea del temporal derecho, trozos de ambos parietales y parte del flanco derecho del hueso frontal, y partes del esfenoides, así como los distintos huesos de la región central de la cara que se halla ligada a un fragmento de la mandíbula izquierda. Se han encontrado también, sueltos, láminas timpánicas del oído derecho y ambos temporales en estado de petrificación.

El tipo de homínido que representa el cráneo suele relacionarse con el complejo industrial achelense. Lo confirma, asimismo, la datación geológica del suelo y los contenidos minerales de los depósitos.

Sin embargo, los útiles del entorno no aportan mucha luz, pues pertenecen a un tipo de industria no determinado. Así, los materiales líticos recogidos *in situ* cuentan un número muy reducido de útiles acabados. De los 270 instrumentos de piedra y hueso animal hallados en la primera capa, sólo veinte estaban terminados. Predominaban los cantos rodados y los martillos de piedra; había seis lascas en total, tres de las cuales eran regulares (dos triangulares y una tercera rectangular) y las tres restantes irregulares. Los núcleos eran amorfos.

Entre los materiales brutos abundaban los de cuarzo y cuarcita. En el manto arenoso que se extiende por la primera capa se encontraron cinco núcleos, cuatro lascas regulares (una con los contornos pulimentados), tres lascas irregulares, una raedera, dos pebbles de cuarzo usados probablemente como martillos de piedra y veintidós lascas gastadas.

Aunque en muy escasa cantidad se descubrieron hachas de mano y cuchillos entre los materiales de la superficie, pero también aquí predominaban los cantos rodados y, en cierta medida, los núcleos.

Pero la ausencia de hachas de mano y de cuchillos entre los útiles hallados *in situ* plantea la cuestión del tipo de industria a que pudiera pertenecer la nueva subespecie de *Homo erectus.* Se han adelantado dos hipótesis para resolver el problema: o bien el material disponible pertenece a un continuum cultural y tecnológico diferente, propio de una industria de fecha posterior o contemporánea a las postreras fases del complejo industrial achelense, o bien pertenece al achelense, y el área excavada muestra una parte de su actividad. La segunda hipótesis le parece más verosímil a Mturi, quien recientemente ha expuesto en *Nature* la estratigrafía y arqueología del hallazgo; se apoya en el descubrimiento de hachas de mano y cuchillos entre los instrumentos de superficie. Más aún, excavando, se ha encontrado otra hacha de mano *in situ.*

Cheops juega al ajedrez

Hace más de 20 años que se viene trabajando en la confección de programas para que los computadores jueguen al ajedrez. A los jugadores de ajedrez humanos les producirá cierta satisfacción saber que, a pesar de ello, los mejores programas no han superado el nivel de los campeonatos de aficionados de grado medio. Los programas de ajedrez llevan nombres tales como Chaos, Ostrich, Treefrog y Eaoin-Shrdlu. El título de campeón de ajedrez entre los computadores lo ostenta en Estados Unidos un programa llamado Ajedrez 4.4, desarrollado en la Northwestern University. El título de campeón mundial lo tiene actualmente Kaissa, un programa escrito en el Instituto de Ciencia del Control, en Moscú.

El optimismo inicial de los programadores de ajedrez se basaba en la premisa de que los computadores serían capaces de examinar todos los movimientos y réplicas posibles con una anticipación de varias jugadas. Y de hecho se han elaborado algoritmos muy ingeniosos para examinar enormes "árboles" de variaciones y para acelerar la búsqueda del movimiento correcto mediante "podas" muy expeditivas. Un computador rápido puede examinar entre 50.000 y 500.000 posiciones durante el plazo típico de tres minutos de que se dispone para hacer una jugada. Por el contrario, un gran maestro humano difícilmente examinará más de 100 posiciones. La habilidad de este último consiste en seleccionar las pocas jugadas que ofrecen más posibilidades, analizándolas con mayor profundidad de la que puede conseguir un computador. La búsqueda del computador no suele ir más allá de cuatro o cinco movimientos y sólo raramente llega a ocho o diez. Esta limitación se debe al crecimiento exponencial del árbol de variaciones a medida que se profundiza la búsqueda. Si a cada movimiento se puede responder con otros diez movimientos verosímiles, la investigación de cinco movimientos implica 100.000 jugadas posibles. Uno de los preparadores del programa Kaissa manifestó a un programador norteamericano: "Si pudiésemos conseguir la investigación de quince movimientos, dominaríamos el mundo".

Puede que el esfuerzo más ambicioso para elevar el nivel del ajedrez de los computadores sea el que están patrocinando conjuntamente el Artificial Intelligence Laboratory y el Laboratory for Computer Science en el Instituto de Tecnología de Massachusetts, bajo la dirección de R. D. Greenblatt. Greenblatt preparó el primero de los programas de ajedrez modernos a finales de los años 60, al que se conoce con el nombre de MacHack. Greemblatt y sus colegas han construido ahora un computador especial, llamado Cheops (abreviatura de *Chess Oriented Processing System*), cuya misión es proporcionar un soporte de "fuerza bruta" para un programa "elaborado". Se entiende por programa elaborado aquel que evalúa la posición fundamental del tablero en cada fase del juego y que dispone de un conjunto de criterios para seleccionar los movimientos más prometedores desde una posición dada. La dificultad estriba en que ni siquiera los grandes maestros pueden ponerse de acuerdo sobre tales criterios y algunos dirían que su número es prácticamente infinito. El papel de Cheops es doble: ha de examinar los movimientos propuestos por el programa elaborado para ver si no implican dificultades no aparentes y ha de considerar por su propia cuenta si una investigación profunda del árbol de posibilidades no puede originar una buena jugada pasada por alto por el programa elaborado. Cheops se preocupa sólo de la ganancia o la pérdida de las piezas, mientras que el programa elaborado es el responsable de toda la estrategia de posiciones.

Cheops está formado por 2200 circuitos integrados que contienen, entre otras cosas, dos conjuntos representativos de los 64 escaques de un tablero de ajedrez: un conjunto "de origen" y un conjunto "de destino". Cada conjunto tiene 64 líneas de salida que indican si un escaque concreto está ocupado o no y por qué pieza. A Cheops se le puede pedir, por ejemplo, que "encienda" todos los escaques a los que puede trasladarse una pieza dada o todos aquellos en los que dicha pieza puede capturar a otra. Como Cheops puede indicar en cada ciclo de operación de 200 nanosegundos la situación en que se encuentran simultáneamente los 64 escaques, puede examinar 150.000 posiciones por segundo.

En consecuencia, en tres minutos puede examinar más de 25 millones de posiciones. Mientras tanto, el programa elaborado, que actúa sobre un computador KL10 de Digital Equipment, estará examinando 40 posiciones por segundo entre variaciones cuidadosamente seleccionadas. El programa completo, que integra el programa elaborado con la "fuerza-bruta" de Cheops, se está escribiendo todavía, pero Greenblatt espera poder hacer una prueba del programa completo frente a jugadores humanos de alto nivel a finales del presente año.

Energía nuclear (1)

De los numerosos puntos de contacto existentes entre las actuaciones de la ciencia y la tecnología, por un lado, y los intereses de la sociedad, por otro, ninguno tiene una actualidad mayor en todos los países del mundo que los problemas planteados por el sector energético y, dentro de él, la utilización de la energía nuclear para fines no militares. La preocupación pública por la programación y puesta en marcha de centrales nucleares para la producción de energía eléctrica hunde sus raíces en causas tan justificadas de temor y ansiedad como los horrores de una guerra termonuclear, el deseo de evitar la progresiva degradación del medio ambiente o los riesgos que para la salud individual y colectiva representaría un nivel de radiación elevado.

Sin embargo, no siempre se consideran adecuadamente en estos debates públicos muchos otros factores implicados, de tipo social, técnico, económico y de actuación a largo plazo, lo que no contribuye a aclarar las cosas ni a permitir que por medio del consenso general se establezcan las soluciones más viables en el momento presente para un problema que es a la vez complicado, importante y urgente. Por medio de esta nota, y las que la seguirán en números próximos, intentamos ofrecer a nuestros lectores un resumen objetivo de los términos en que actualmente está planteada la cuestión. Servirán de marco de referencia para los estudios, análisis e informaciones sobre aspectos concretos que *INVESTIGACION Y CIENCIA* irá insertando en el futuro.

Por lo pronto, conviene dejar bien establecido que el recurso a la utilización de la energía nuclear se plantea como una necesidad y no como una mera consecuencia del progreso científico o tecnológico. No se trata tanto de que el nivel actual de los conocimientos teóricos y aplicados permita optar entre la continuación del sistema actual (en el que las necesidades energéticas se cubren por medio del petróleo y, en menor medida, del gas natural, el carbón o la energía hidroeléctrica) y otra alternativa proporcionada por la energía nuclear, que para competir con las otras fuentes tendría que ofrecer claras ventajas en términos de coste, disponibilidad a largo plazo, facilidad de obtención y ausencia de mayores efectos nocivos secundarios. Lo que hace ineludible el planteamiento del tema es la previsión de que, no sólo a largo plazo sino para un futuro muy inmediato, el balance energético de cada uno de los países, especialmente

los desarrollados, y, en términos más generales, el balance mundial, muestra un claro desequilibrio, ya que las fuentes de energía actualmente disponibles no serán suficientes para satisfacer las necesidades entonces existentes.

Una somera consideración histórica ayudará a ver más claramente cómo ha evolucionado la demanda de energía y cómo se la ha satisfecho.

Hasta finales del siglo XVIII las exigencias energéticas de la humanidad fueron mínimas y quedaban cubiertas por la energía animal y por combustibles naturales como la madera o las grasas animales y vegetales, seguidos muy de lejos por el carbón o el petróleo superficiales y sus derivados. El siglo XIX representa el primer punto de inflexión de la curva de la demanda energética, empezándose a explotar más intensamente los yacimientos de carbón para su utilización en metalurgia y en obtención de gas de alumbrado. Pero hasta el comienzo del último cuarto del siglo XIX no existen las condiciones tecnológicas para la utilización de grandes cantidades de energía. Es en este período cuando las máquinas de vapor empiezan a ser utilizadas ampliamente y, sobre todo, cuando la energía eléctrica empieza a encontrar aplicaciones prácticas para el alumbrado y la generación de movimiento. Aproximadamente de la misma época data el comienzo de la explotación del petróleo (se considera que el primer pozo perforado a la manera actual lo fue en 1859), que se utilizaba fundamentalmente para fines de iluminación. La energía eléctrica, y luego el petróleo y sus derivados, especialmente la gasolina, presentaban unas características ideales para que su consumo se generalizase: alto poder energético y facilidad de distribución. Sin embargo, el crecimiento del consumo fue muy paulatino hasta bien entrado el siglo XX.

Hacia 1950 es cuando se produce el segundo y decisivo punto de inflexión en la curva del consumo. Entre 1950 y 1975 el consumo energético mundial aumenta cerca de un 400 %, mostrando una tendencia sostenida e incluso creciente, si bien a partir de 1973 las perspectivas cambiaron bastante. En efecto, diversos factores hacen que en este año se produzca un hecho decisivo que conmueve profundamente la economía mundial y frena el ritmo de desarrollo de todos los países: el incremento sin precedentes en el precio del crudo de petróleo. Mientras que a lo largo del período 1950-1973 el precio del petróleo, en moneda constante, había ido disminuyendo progresivamente, en otoño de 1973 se multiplica por cuatro. El pilar básico del desarrollo experimentado en esta época fue la disponibilidad ilimitada de energía barata, en forma de petróleo. A partir de este momento ya no es ni tan barata ni, desde luego, ilimitada.

Dentro de la incertidumbre que toda estimación comporta, hay un acuerdo suficientemente generalizado de que, si el ritmo de consumo mantiene la tónica de incremento del período 1950-1975, antes del año 2000 se alcanzaría el límite máximo de crecimiento de la oferta de petróleo y, a partir de ese momento, su disponibilidad disminuiría bruscamente, a pesar de que durante todos estos años el esfuerzo de explotación de nuevos yacimientos hubiese continuado con la máxima intensidad. El petróleo se convertiría, pues, en un producto de escasez creciente y cuyo precio aumentaría en forma proporcional. En el año 2100 estaría prácticamente agotado. El año 1995 o el año 2000 están demasiado próximos para que no sea urgente planificar la forma en que va a hacerse frente a esta amenazadora contingencia, sobre todo si se tiene en cuenta que el petróleo proporciona entre el 50 y el 80 % del total de la energía consumida.

Evidentemente, podría suceder que el ritmo de expansión de los países industrializados, que incidentalmente consumen entre el 70 y el 80 % de toda la energía mundial, disminuyese, precisamente como consecuencia de la crisis energética, pero es de suponer que el esfuerzo de los países subdesarrollados por salir de esta situación supondría una compensación de sentido contrario, aunque también esté claro que el desarrollo de estos últimos se verá muy dificultado por la no disponibilidad de energía a precios baratos. Quiere decir todo esto que, con independencia de que los hechos se produzcan unos años antes o después de lo que indican las estimaciones actuales, todos los países deben prever la forma en que obtendrán la energía que necesitan de fuentes distintas del petróleo. El problema es especialmente grave para aquellos que no disponen de producción propia y dependen fuertemente del petróleo importado (situación en la que se encuentran muchos del mundo y todos los de Europa occidental), por el grave desequilibrio que esta masiva importación origina en sus balanzas de pagos.

En este contexto aparece la energía nuclear como una de las posibles vías de solución. Existen otras, y la evaluación relativa de lo que cada una de ellas pueda aportar es una de las tareas más importantes y más difíciles de la situación presente. Por consiguiente, en el próximo número revisaremos someramente sus características y posibilidades.

Diabetes: ¿una sola enfermedad?

Desde hace cincuenta años aproximadamente –desde el descubrimiento de la insulina por Frederick Banting y Charles Best a principios de la década de los veinte– se ha venido considerando la diabetes como una enfermedad bastante bien conocida y, hasta cierto punto, controlable. Se la cree producida por una deficiencia, al parecer congénita, de las células del páncreas encargadas de la secreción de insulina, que es la hormona responsable de la regulación del metabolismo normal de glúcidos en el cuerpo. Los síntomas químicos que dan cuenta de este trastorno metabólico debilitante son perfectamente conocidos: acumulación de azúcar (glucosa) en el torrente sanguíneo, excretándose por vía urinaria el azúcar sobrante no metabolizado. Mediante la simple estratagema de inyectar diariamente al diabético una dosis de insulina y someterlo a una dieta rigurosa, se consigue estabilizar los casos graves y evitar sus complicaciones y el coma diabético.

A pesar de lo cual, y según pone de manifiesto una obra de reciente aparición, la diabetes "sigue siendo uno de los enigmas más obscuros de la medicina. Nadie sabe cuáles son sus causas, cómo prevenirla, si es una o varias a la vez o cómo atajar los efectos profusamente devastadores que produce en el organismo" (Bertrand E. Lowenstein y Paul D. Preger: *Diabetes: New Look at an Old Problem,* Harper and Row, 1976).

Los índices estadísticos sobre la incidencia de la diabetes no son precisamente alentadores. Según el Public Health Service, de Estados Unidos, es la causa más destacada de la ceguera, de las disfunciones renales, de los trastornos nerviosos, de los desórdenes circulatorios, de la gangrena de las extremidades y de diferentes disfunciones sexuales. También se ha relacionado la diabetes con el infarto, la apoplejía y otras enfermedades cardiovasculares. En resumen, cincuenta años de aplicación de insulina sólo han conseguido que el promedio de vida del diabético sea todavía un 25 % inferior que el del no diabético. Por poner las cosas más obscuras, descubrimientos recientes hacen sospechar que los medicamentos por vía oral prescritos para regular el nivel de

azúcar en sangre no producen efectos durables en aquellos pacientes cuyo grado de diabetes no exige un tratamiento insulínico.

Ante este cuadro de perfiles negativos, surge una renovada investigación sobre la incidencia, causas y prevención de la diabetes. Lo que es más importante, algunos replanteamientos del problema, todavía en fase preliminar de investigación, podrían mejorar las esperanzas del diabético. Por una parte, según comentaba recientemente el *British Medical Journal,* "se está llegando al convencimiento de que la diabetes no es una entidad única, sino que responde a un complejo de varias enfermedades cuyo común denominador es el elevado nivel de azúcar en sangre". La forma más extendida de ese mal, la denominada diabetes de la edad madura (puesto que suele aparecer pasados los cuarenta años), evidentemente "manifiesta una fuerte tendencia hereditaria", dice el citado editorial, que sigue: "Se ha calculado que el 60 % de los hijos de padre y madre diabéticos en su edad madura padecerán a su vez diabetes benigna hacia los 60 años. Por otra parte, contrariamente a lo que se venía sosteniendo, la diabetes clásica contraída en la juventud parece tener un componente genético mucho más débil. Lo cual resulta alentador para esa clase de pacientes, que se preguntan por las posibilidades que tienen sus descendientes de contraer su propia enfermedad, y a quienes se les puede asegurar que la probabilidad del riesgo que sus hijos corren no es quizá mayor del 1 %.

Las pruebas de un origen no hereditario de diabetes aguda en la juventud, o de diabetes que requiere un tratamiento por insulina, proceden de fuentes dispares. Al haberse comprobado que la incidencia en la juventud varía con la edad, estación, año y zona geográfica, se supone que tal vez los factores ambientales desempeñen una función importante. Investigaciones recientes en Gran Bretaña y Estados Unidos han despertado gran interés ante la posibilidad de que algunos virus, tal el virus de las paperas o el virus Coxsackie B4, sean agentes desencadenantes de diabetes en los jóvenes. (Ya se ha demostrado por vía experimental que los virus provocan la diabetes en ciertos animales.) A fin de descubrir la posible conexión entre virus y diabetes, los científicos británicos han realizado una serie de pruebas con hermanos gemelos idénticos, en que algunos grupos estudiados eran discordantes (sólo un gemelo padecía diabetes) y otros concordantes (ambos gemelos idénticos la poseían). Si bien las investigaciones de última hora manifiestan que "no existe una prueba directa" de diabetes por inducción vírica en los jóvenes gemelos, esos mismos investigadores informan también que "la edad y el tiempo en que se produjo la enfermedad sugieren que los factores ambientales influyen decisivamente en la diabetes en gemelos".

Refiriéndose a los estudios llevados a cabo con gemelos idénticos en Gran Bretaña, Lowenstein y Preger apostillan en su obra citada: "Aun cuando las pruebas de que se dispone no bastan para certificar el carácter determinante del virus en la diabetes, sí alienta a los investigadores a seguir por ese camino, porque si realmente se produjera la diabetes por acción de un virus algún día se llegaría a la vacuna contra el mismo". Y, tomando en consideración no sólo las investigaciones sobre etiología vírica sino otras líneas de estudio también (verbigracia, el trasplante total o parcial de un páncreas sano o la implantación de algún aparato artificial que provoque la secreción de insulina), Lowenstein y Preger concluyen que "hay motivos para esperar que dentro de cinco años la diabetes podrá controlarse, seguramente, mucho mejor".

Liberación femenina

Casi todos los aspectos de la vida moderna han sufrido la influencia revolucionaria de la tecnología, sin que la vida en el hogar sea una excepción. La introducción en él de las máquinas alteró, en concreto, el papel del ama de casa, pero no siempre en la forma que podía esperarse. Ruth S. Cowan, de la Universidad del Estado de Nueva York en Stony Brook, ha analizado en la revista *Technology and Culture* el desarrollo de la tecnología doméstica y su influencia sobre las mujeres de la clase media norteamericana en el período comprendido entre las dos guerras mundiales.

La interpretación tradicional de la influencia de la tecnología sobre la familia moderna es la siguiente: con anterioridad a la industrialización, la familia era la unidad social básica, que producía y elaboraba prácticamente todo lo que necesitaba para su propia sustentación. En las familias pre-industriales las mujeres tenían mucho trabajo y casi todo su tiempo resultaba ocupado por las tareas domésticas. En una sociedad industrializada la unidad familiar ya no es el centro fundamental de producción y, a medida que su función va disminuyendo, se van haciendo menos rígidos los vínculos que la mantenían. En estas familias las mujeres tienen poco que hacer y muchas sienten ansiedad sobre cuál sea su papel.

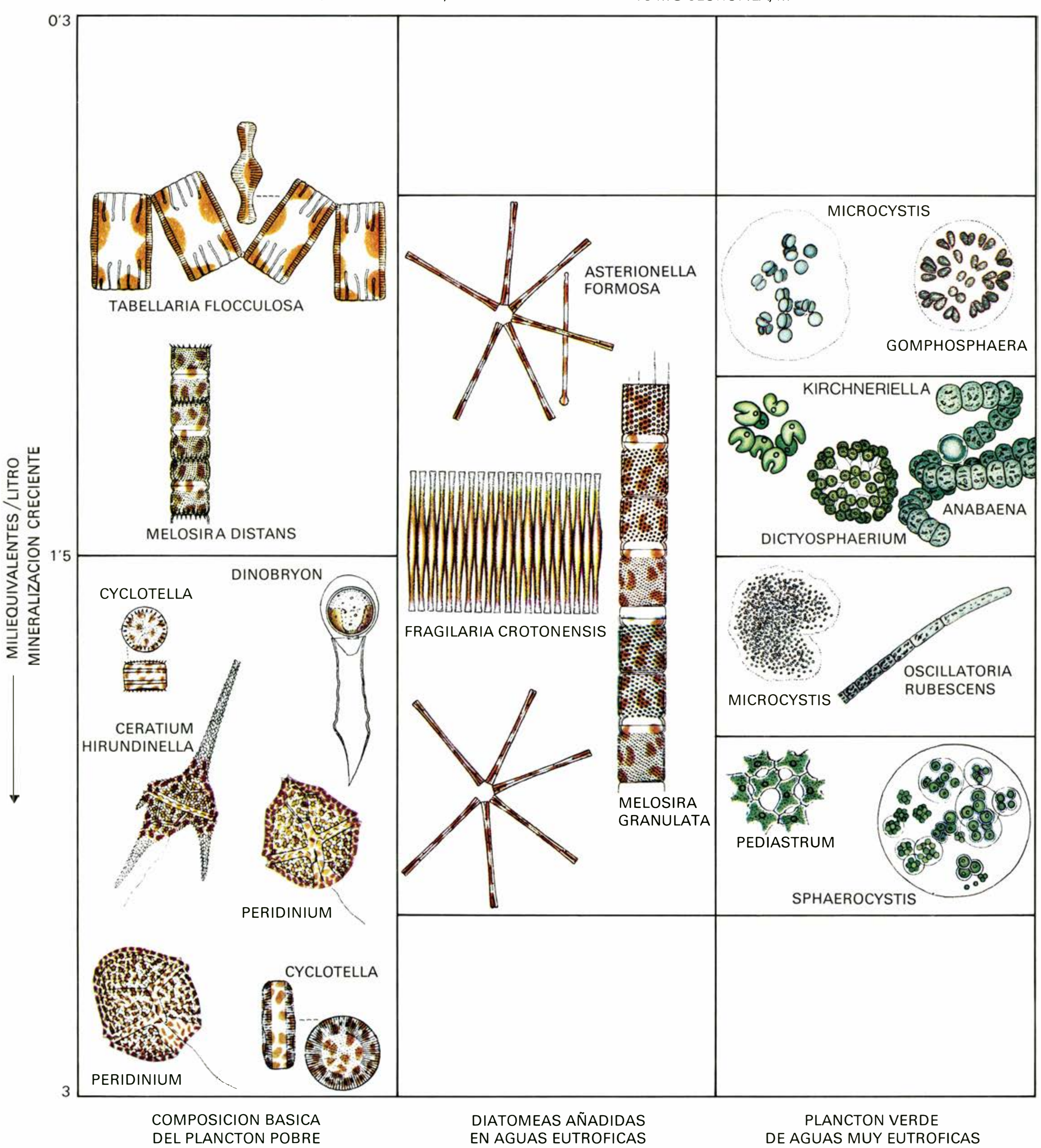
Cowan ha examinado familias pre y post-industriales, a través de la forma en que quedaban reflejadas en los artículos y en los anuncios de las revistas femeninas de los años 20 y 30, y ha llegado a la conclusión de que la interpretación tradicional es inadecuada. Aunque el ama de casa de clase media tenía mucho trabajo para ocupar su tiempo en la familia pre-industrial, solía tener también ayuda doméstica, pagada o gratuita. A medida que se fueron introduciendo en el hogar los aparatos eléctricos, tales como planchas, cocinas y lavadoras, el ama de casa pasó a realizar por sí misma una parte creciente de este trabajo. Y en este proceso quedó reducida del papel de un director de empleados al de un trabajador doméstico.

El ama de casa de clase media tenía que pasar más tiempo con sus hijos, tenía que cocinar la mayor parte de las comidas y normalmente tenía que hacer casi todas las compras de alimentos y vestidos de la familia. Dado que estas tareas resultan más fáciles con una cierta capacitación, se esperaba de ella que se preocupase por adquirirla y que dominase un amplio abanico de tareas que previamente no eran de su competencia o que ni siquiera existían.

Conforme cambiaba el puesto de ama de casa, iba cambiando también la ideología relacionada con él.

Antes de la guerra, el hecho de tener que hacer las tareas domésticas era como una carga que había de soportarse hasta que se encontrase un sirviente cualificado. Después de la guerra, ya no parece una carga, sino una expresión de amor del ama de casa por su familia y una obligación que no puede confiarse a un sirviente.

Cuando las sociedades se industrializan, la estructura de la fuerza laboral sufre un proceso de gran diferenciación en el que los trabajadores individuales se hacen más especializados y desaparece el contexto emocional del trabajo. Según la interpretación tradicional del efecto de la tecnología sobre la familia, deberían observarse las mismas consecuencias en el hogar, pero lo que se observa es exactamente lo contrario. "La industrialización del hogar", concluye Cowan, "fue un proceso muy diferente de la industrialización de otros medios de producción y el resultado de dicho proceso no fue ni el que se nos ha hecho creer que fue ni el que los estudiosos de las otras revoluciones industriales pudieran predecir".



ORGANISMOS DEL FITOPLANCTON, frecuentes en los embalses. Se pueden utilizar como indicadores de distintas combinaciones de características ambientales. Muy aumentados, *Tabellaria* y *Melosira distans* son exclusivas de los embalses de aguas poco mineralizadas, mientras que *Ceratium hirundinella, Dinobryon* y *Cyclotella* son especialmente abundantes en los de aguas alcalinas, ricas en calcio. Cuando el grado de eutrofia es mayor, se añaden densas poblaciones de los tres géneros de diatomeas representados en el centro. Cianofíceas y clorofíceas *(a la derecha)* proliferan al aumentar la eutrofia, según combinaciones que dependen de las restantes propiedades del agua.

Biología de los embalses

La construcción de un gran número de embalses en un país sin lagos constituye una situación experimental única para el estudio del funcionamiento de éstos y de la colonización de nuevos ambientes

Ramón Margalef

Existe una progresiva preocupación mundial por la calidad de las aguas continentales. Además del problema de la contaminación creciente, está el de la fertilización paulatina de las aguas naturales, que van aumentando su producción en materia orgánica, con la consiguiente pérdida de calidad. Por si fuera poco, las necesidades de agua crecen constantemente, teniendo que destinar nuevas reservas a usos domésticos. No sólo preocupa la fertilización de las masas de agua que se están utilizando actualmente, sino también la extensión del fenómeno a aguas que pueden emplearse el día de mañana, porque, cuando llegue el momento de servirse de ellas, su calidad estará ya muy disminuida. Entre las aplicaciones prácticas más importantes de la limnología, ciencia que estudia la ecología de las aguas dulces, está el detectar y predecir la fertilización y proponer remedios. En nuestros días, el uso del agua de ríos y embalses para la refrigeración de unidades productoras de energía, térmicas y nucleares, ha añadido otro argumento a la preocupación por la conservación, regulación o gobierno de las características del potencial hidráulico de un país.

España ha pasado de ser un país sin lagos a otro con más de 700 embalses. Los únicos lagos naturales eran el de Sanabria, en el NW, y el de Bañolas, en el NE, glacial el primero y cárstico el segundo. A los embalses se aplica a veces la designación de lagos artificiales; aunque no son verdaderos lagos, representan superficies de agua estancada, reciben afluentes y su caudal se destina a usos diversos, entre otros, el suministro de agua a ciudades. Al igual que los lagos naturales, están amenazados por la contaminación y la fertilización. Al ser ésta, en su mayor parte, un problema unido a la acción del hombre, se podría decir que los embalses, nacidos como una consecuencia de la civilización, ya desde el principio están fatalmente condenados a la fertilización, o a la eutrofización, su sinónimo técnico, que a veces aparece y aumenta en ellos con suma rapidez. A la fertilización hay que agregar la acumulación de sedimentos en el fondo del embalse, lo cual reduce su capacidad. La acumulación de nutrientes solubles y de sedimentos insolubles aparece frecuentemente asociada, al ser ambos consecuencia de una deficiente conservación de los suelos en la cuenca del embalse.

Además de la fertilización o eutrofización y del relleno de sedimentos, los embalses plantean otros problemas prácticos: posibilidad de extraer un beneficio a través de la pesca y de otras formas de explotación, verbigracia, recreo y turismo, lo que exige medidas de conservación y corrección. Resulta asimismo interesante la atención a la calidad del agua, su composición, que puede regularse eligiendo debidamente la profundidad de extracción en cada embalse, o con un régimen apropiado de gobierno cuando existen varios embalses unidos en cadena. Finalmente, la composición del agua o los materiales sólidos que contiene poseen implicaciones sanitarias, no sólo por poder albergar gérmenes patógenos, sino por la presencia y concentración de determinados elementos o compuestos químicos indeseables y que, habitualmente, son el resultado de la actividad humana.

Si no echamos en olvido la industria de los castores, debemos afirmar que la represa de los ríos es una actividad zoológica prehumana. Nunca, sin embargo, dicha actividad había tenido la importancia que hoy ha asumido ni había afectado a una superficie tan extensa del planeta. El número de embalses es muy elevado, y su construcción, lejos de detenerse, crece a un ritmo acelerado. En todo el mundo, en el año 1972, había, por lo menos, 40 embalses de más de 1000 km² y 260 de superficie comprendida entre 100 y 1000 km². Paralelamente a su construcción, se ha ido incrementando el interés científico por ellos. En la Unión Soviética existe el mayor instituto dedicado exclusivamente a la investigación de los pantanos, a orillas del embalse de Rybinsk, en el Volga. Los gigantescos embalses africanos, en los ríos Zambezi (Kariba), Volta (Akosombo) y Nilo (Asuán) han dado origen a problemas muy variados y han estimulado el consiguiente interés de la ciencia por resolverlos. En casi todos los países europeos y en Estados Unidos se han hecho numerosos estudios de sus embalses respectivos, que en algunos casos han alcanzado gran detalle. El interés ecológico por los embalses ha de ser forzosamente mayor en un país desprovisto de lagos, cual es España, que sólo dispone de dos lagos naturales de más de 1 km en su distancia máxima y más de 40 metros de profundidad: los citados más arriba de Sanabria y Bañolas. La construcción de más de 600 embalses grandes, en el curso de unas décadas, en un país tan falto de aguas continentales, constituye, ciertamente, un experimento ecológico único que convenía no desaprovechar. Nos hallamos ante un número muy grande de nuevos ambientes de agua dulce, establecidos en las condiciones más diversas, que se prestan excepcionalmente bien al estudio

de los procesos de colonización y organización de nuevas comunidades de organismos. La limnología, que como hemos dicho se ocupa de tales procesos en los lagos, encuentra en los embalses unas condiciones experimentales únicas, de cuyo estudio habrán de beneficiarse la gestión de los mismos y la construcción de otros nuevos.

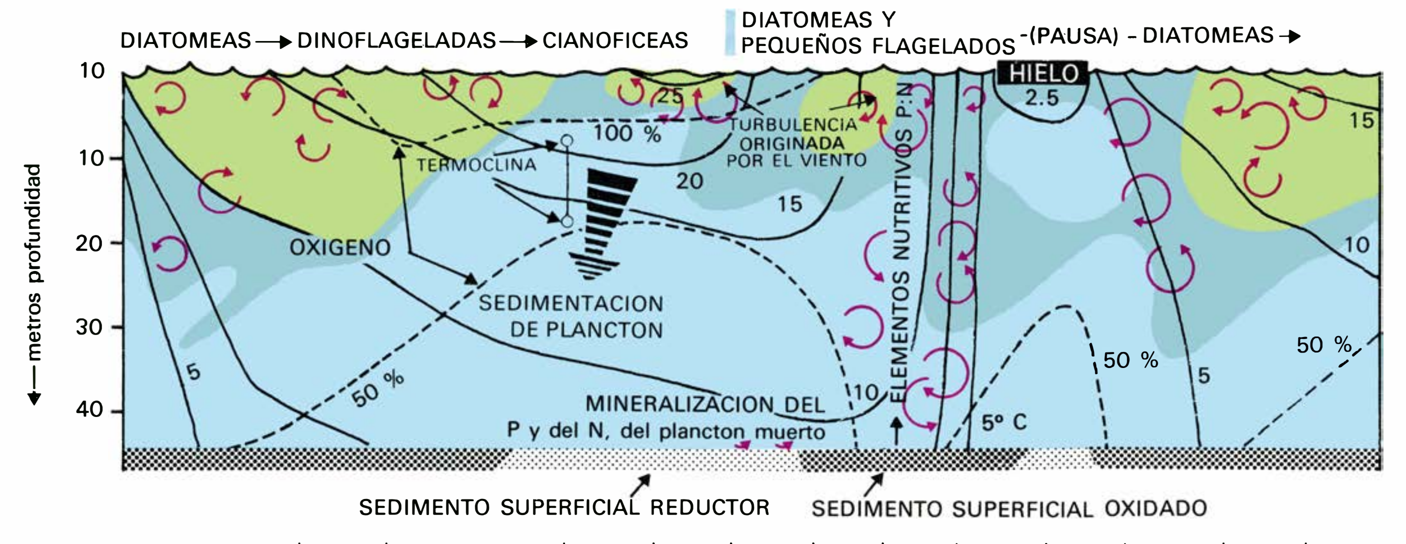
Desde el punto de vista biológico, un embalse equivale, hasta cierto punto, a un lago artificial, y puede estudiarse por comparación con un lago natural. Los lagos naturales que ocupan un valle cerrado por una colada de lava o por una morrena glaciar no difieren sustancialmente de los embalses. Tal ocurre, por ejemplo, con los lagos marginales alpinos. Pero entre un embalse y un lago marginal hay un motivo de diferencia significativo: la edad. Los lagos marginales alpinos adoptaron su forma actual y se poblaron hace varios millares de años, pero no más de 11000, que es el tiempo transcurrido desde las postrimerías del último período frío; la edad de un embalse español típico apenas si sobrepasará el medio siglo, porque el acúmulo de sedimentos, por erosión de las laderas bajo influencia humana, es muy rápido y lo rellenarán.

Hay que precisar más en la comparación entre pantano y lago natural. En aquél, el flujo del agua suele ser más intenso, variable e irregular que en éste. Con pocas reservas podríamos definir el embalse como el híbrido de río y lago; como en los ríos, predomina hacia su cola el transporte horizontal; y, como en los lagos, se manifiesta, junto a la presa, la estratificación gobernada por la dirección de la gravedad y de la luz.

De las múltiples alteraciones que el hombre introduce en la naturaleza, la construcción de embalses es de las menos criticables, porque retarda el flujo del agua y estabiliza así, relativamente, las comunidades de organismos asociados al agua o dependientes de ella. La naturaleza, sin embargo, no tiende a formar embalses, sino a complicar el curso de los ríos, dando meandros que luego se cortan y producen un cortejo de lagunas, donde prosigue la evolución de la vida acuática. La construcción del embalse por obra del hombre puede considerarse como un substituto de ese proceso natural, en el sentido de que ofrece oportunidades para la colonización y desarrollo de comunidades de organismos, efímeras a una escala geológica del tiempo.

Las razones que inducen a la construcción de embalses son múltiples. Sólo cuando su función principal es regularizar los flujos y evitar inundaciones puede decirse que van a favor del proceso natural de la sucesión y organización creciente de los ecosistemas. Cuando se destinan a la producción de energía, riego, consumo doméstico y usos industriales, la forma como el hombre gobierna su flujo no suele coincidir con la más conveniente para la llamada conservación de la naturaleza; lo cual se pone de manifiesto en el impacto que suele sufrir la vida acuática y la de las riberas con las fluctuaciones de nivel del embalse. Estas provocan que la capacidad de producción se desvíe de la superficie del suelo o del sedimento a las aguas libres; en efecto, por las variaciones de nivel, zonas extensas de la orilla y, sobre todo, en la cola del embalse quedan inundadas o secas, alternativamente y, además, se incrementa la productividad del plancton o vida suspendida en las aguas, en detrimento de los organismos del bentos, o fondo, y litorales. El agua adquiere entonces un matiz verdoso. Esta y otras consecuencias del régimen a que están sometidos los embalses pueden carecer de importancia práctica cuando las aguas se destinan al riego o a la producción de energía eléctrica, objetivos contra los cuales atentan gravemente la erosión de las laderas y el relleno de sedimentos sólidos, que reducen la capacidad del embalse. Mas para otros usos, habrá que preocuparse también por las alteraciones producidas por la afluencia al pantano de substancias solubles procedentes de la misma cuenca, cuando dichas substancias regulan la biología del embalse. Su influencia sobre la vida del embalse, en el sentido de favorecer más el plancton (vida suspendida) o al bentos (vida del fondo), depende también, según se dijo, del régimen o gestión de la presa.

Las alteraciones más notorias en las características ecológicas consisten en la fertilización y enverdecimiento del agua y en el agotamiento del oxígeno en las capas profundas. Son dos procesos complementarios que reciben el nombre común de *eutrofización.* Si el caudal se des-

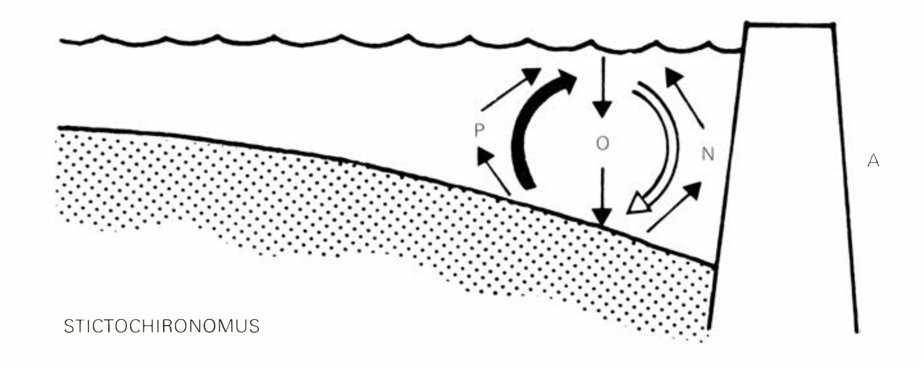


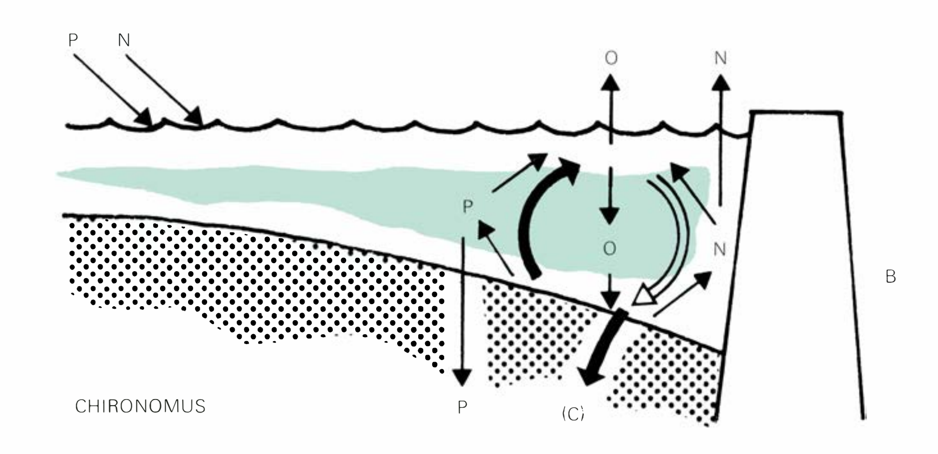
DINAMICA DEL CICLO ANUAL de un lago de la zona templada. Las profundidades en ordenadas y el tiempo en abscisas. La temperatura se representa por isotermas de trazo continuo, y la concentración de oxígeno (en % de saturación), por trazos discontinuos; las flechas indican la turbulencia del agua originada por el viento; la diferente intensidad de color simboliza el correlativo desarrollo del fitoplancton. Las características de la superficie del sedimento varían según el contenido del oxígeno del agua del fondo, siendo reductor en épocas de anoxia y oxidado cuando hay oxígeno. El ciclo de los organismos en el fitoplancton, en razón de la época, se indica en la parte superior. En otoño, la mezcla de las aguas permite reanudar el ciclo. La pausa explica la ausencia de organismos al helarse el medio.

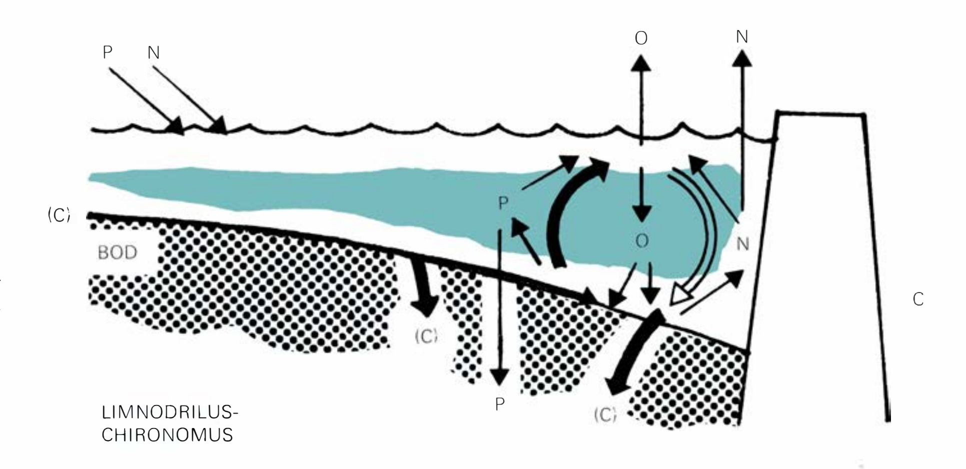
tina a riego o salto de agua para producir electricidad, es indiferente la cantidad de plancton que arrastra el agua, o si el contenido de oxígeno en las capas profundas se agota o no durante el verano. Pero si se trata de abastecer con su reserva a las ciudades, destino éste que irá intensificándose en el futuro, se hubieran construido los embalses para esa finalidad o para otra distinta, la eutrofización del agua sí que importa, y mucho. Y conviene limitar en lo posible dicho proceso de eutrofización, entre otras razones, porque una de sus secuelas es la acumulación de nutrientes en el sedimento, nutrientes que pueden ir retornando paulatinamente al agua, aun cuando, para entonces, se hayan adoptado medidas de protección. Si desde ahora se deja que el proceso de eutrofización siga su curso, hipotecamos el futuro del embalse para un posible consumo humano.

Con ser grave, la eutrofización no es el único problema que plantea el estudio limnológico de una presa. El uso de su caudal para la refrigeración de centrales productoras de energía perturba el balance térmico y los movimientos del agua, afectando por tanto a la vida que se desarrolla en ella. Y tal vez sea más grave todavía la incidencia cuando se trate de la refrigeración de centrales nucleares por el posible riesgo de que dosis, aunque sean pequeñas, de radiactividad se acumulen en el sedimento y se movilicen luego, en virtud de un mecanismo substancialmente idéntico al que explicaremos a propósito del fósforo y de varios metales. Intentaremos plantear adecuadamente el alcance de estos problemas y sus líneas de solución.

Uno de los enfoques tradicionales de la ciencia limnológica es la limnología regional, que pretende una tipificación de los lagos compatible con su distribución geográfica. Esta metodología se desarrolló a principios de siglo en Escandinavia y norte de Alemania, por Naumann y Thienemann. De forma esquemática, conduce a distinguir dos tipos opuestos de lagos que, en una versión más modernizada, llegan a considerarse como los extremos de un eje de variación continua, que corresponden, respectivamente, a los lagos oligotróficos y a los eutróficos. Los lagos oligotróficos son profundos, excavados en roca poco soluble, de aguas transparentes y pobres en elementos químicos (principalmente el fósforo y el nitrógeno); el plancton escasea y el agua permanece bien oxigenada en profundidad durante todo el verano. Los lagos eutróficos típicos, a que se referían los primeros limnólogos, eran lagos







CICLO DE ELEMENTOS LIMITANTES en distintos tipos de embalses. El transporte neto hecho en el cuerpo de los organismos o de sus restos se simboliza con la flecha blanca: el retorno en forma de substancias disueltas, con la flecha negra. En un embalse teóricamente oligotrófico (A), el ciclo quedaría cerrado dentro del agua, con un intercambio mínimo con la atmósfera o con el sedimento. En el caso general de embalse eutrófico (B) existen ciclos secundarios que se extienden hacia la atmósfera (oxígeno y nitrógeno) y hacia el sedimento (materia orgánica, fósforo y metales). En el embalse eutrófico que recibe materia orgánica (C), la situación se agrava por el consumo excesivo de oxígeno (BOD). El color verde alude a la concentración de plancton, el negro a la presencia de materia orgánica en el sedimento. Aquí se incorporan los organismos más representativos de los distintos tipos de embalse.

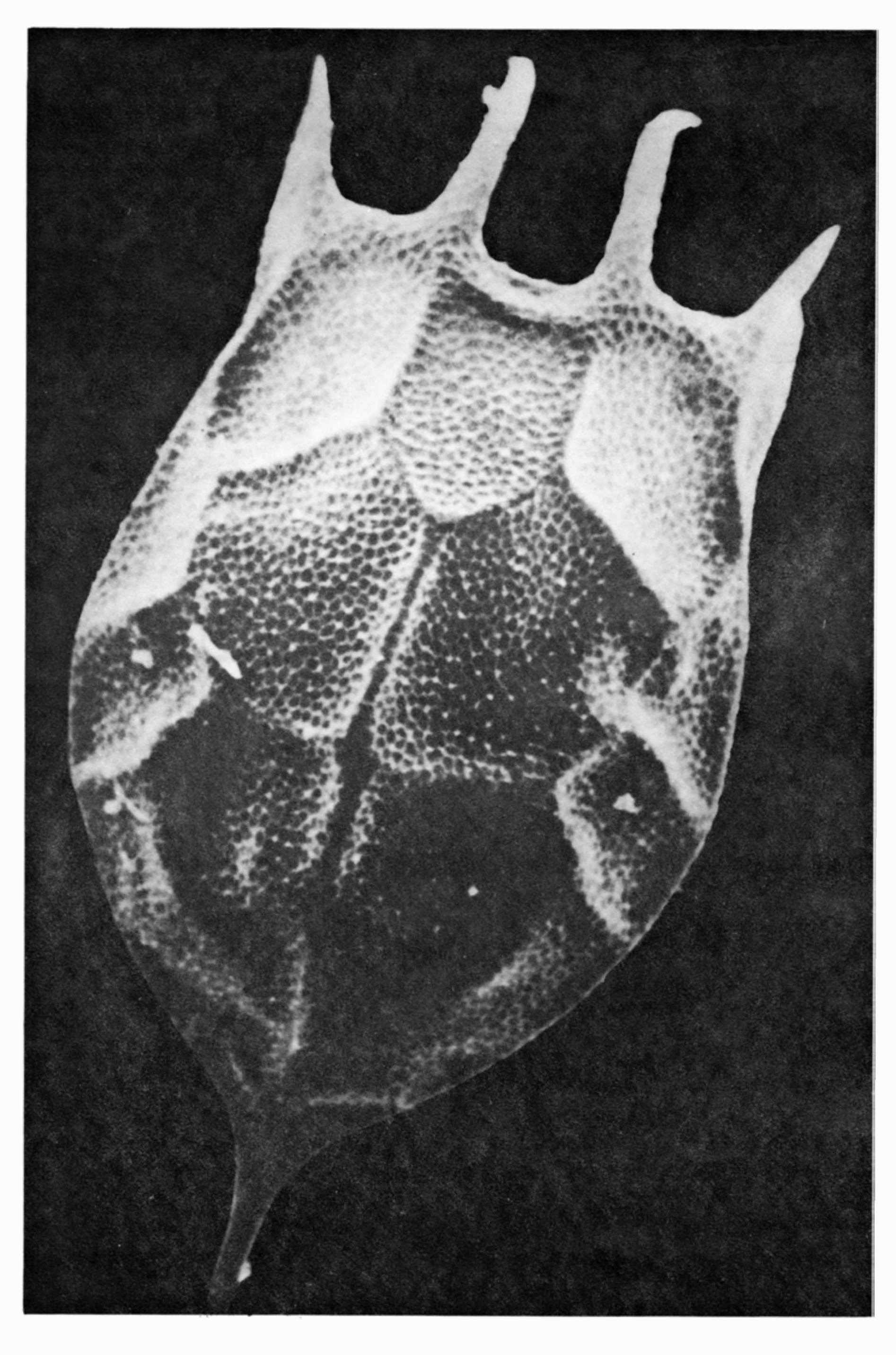
de llanuras cubiertas de sedimentos glaciales, relativamente ricos en materias solubles y en los nutrientes necesarios, a veces como consecuencia de la explotación agrícola de la cuenca; sus aguas aparecían poco transparentes, verdosas y ricas en plancton y agotaban su oxígeno en profundidad hacia finales de verano, si no antes.

El interés por la clasificación de los lagos según el eje oligotrofia-eutrofia aumentó al comprobar que el relleno

parcial por sedimentos y la aportación de elementos nutritivos (fósforo, sobre todo) o de materia orgánica, a consecuencia de la acción humana, podía determinar la evolución desde una situación oligotrófica, o escasamente eutrófica, a otra eutrófica, o plenamente eutrófica. Semejante alteración no podía pasar inadvertida, pues se extendía como una plaga; afectaba a todos los lagos de los países desarrollados, dando origen a la voz eutrofización, ya generalizada en su uso, para designarla. Eutrofización significa, pues, dificultad creciente para obtener agua de buena calidad, toda vez que las aguas profundas de los lagos oligotróficos, ricas en oxígeno, se van degradando; al eutrofizarse, el agua es de calidad baja, irregular, pobre en oxígeno, rica en materia orgánica, llega a despedir malos olores y ataca al hierro y al cemento.

El proceso de eutrofización varía en razón de la zona climática en que se encuentre la presa. En los lagos y embalses de regiones templadas, la eutrofización general y la cultural en particular se explican fácilmente. En los lagos tropicales, cuyas aguas pueden estar permanentemente estratificadas según gradientes de densidad, ocurren fenómenos de otro tipo, de los que ahora no podemos ocuparnos. Por el contrario, en la zona templada, la temperatura de las masas de agua continentales se uniformiza verticalmente durante la estación fría, o cuando la temperatura superficial desciende tanto que el agua se hiela; existen entonces dos períodos, a saber, antes y después de la formación del hielo, en que la temperatura es uniforme de arriba abajo. En esos momentos puede producirse la mezcla vertical ya que el efecto del viento encuentra muy poca resistencia al no oponérsele una estratificación vertical por densidades. En dichas ocasiones no sólo desaparece la variación en la temperatura (el gradiente térmico), sino que se consigue, además, uniformidad en todas las propiedades, incluidas las concentraciones de oxígeno y de compuestos de nitrógeno y fósforo. Los compuestos inorgánicos de carbono (carbónico y bicarbonatos) son sobradamente suficientes para las necesidades de los organismos en todo tiempo, por cuya razón no merece la pena ocuparse de ellos ya que no regulan la cantidad de vida producida.

El ciclo de la producción biológica se ajusta estrechamente a esos cambios ambientales. Así, la mezcla vertical de la estación fría dispersa a muchos organismos del plancton vegetal hacia abajo y al fondo, lejos de la luz, dispersión que implica pérdidas y consiguiente freno en el desarrollo de sus poblaciones. En primavera y verano, el agua se calienta en la superficie y el lago o embalse se estratifica, es decir, se origina un gradiente térmico disponiéndose las capas más cálidas, y más ligeras por tanto, sobre las más frías y más densas del fondo. Como consecuencia de la agitación del agua superficial por la acción del viento, el gradiente térmico vertical del verano no es progresivo, sino que las capas superficiales están muy mezcladas y tienen una temperatura uniforme. La temperatura desciende de una forma relativamente brusca (un grado por metro, o más) a nivel de la termoclina, que separa la zona superficial calentada por el sol de la zona profunda fría. La termoclina suele formarse entre los 5 y los 20 metros de profundidad, según los casos; los estratos situados por encima de ella constituyen el epilimnion, y suelen coincidir con la masa de agua que recibe iluminación suficiente para que en ella sea posible la fotosíntesis por los organismos del plancton. Lo cierto es que en primavera y verano se desarrolla el fito-



KERATELLA COCHLEARIS, **rotífero muy común en las aguas eutróficas y muy frecuente en las menos mineralizadas (Servicio de Microscopia electrónica de la Universidad de Barcelona).**

plancton y se acumula rápidamente en el epilimnion. La cantidad de plancton vegetal, o fitoplancton, que se produce es proporcional a la cantidad de nutrientes disponibles, más los que llegan con los afluentes, procedentes del lavado de los suelos de la cuenca y, por último, con las aguas residuales. Si se produce mucha materia vegetal en el lago o en el embalse, el agua se sobresatura de oxígeno en las capas superficiales; parte de este oxígeno que resulta de la fotosíntesis escapa, por difusión, a la atmósfera, a favor del gradiente natural en su concentración, que va de más a menos.

Parte del fitoplancton sedimenta y parte es devorado por el zooplancton. Los animales del plancton, en el mar y en àguas dulces, más en aquél que en éstas, tienen la costumbre de comer en las capas superficiales y de excretar en las profundas, por lo que contribuyen al transporte vertical hacia abajo de la materia orgánica formada en las capas altas. Esta materia orgánica, pues, se oxidará y descompondrá por debajo de la termoclina, en el hipolimnion. Así, en el lago estratificado, el epilimnion es la zona de producción de materia orgánica y el hipolimnion de descomposición y retorno de los elementos al medio, en forma inorgánica. Cuanto más productivo, más eutrófico, sea un lago, más oxígeno necesitará para oxidar la materia orgánica sintetizada; por consiguiente, el oxígeno del hipolimnion irá consumiéndose hasta su total agotamiento.

El proceso de eutrofización comporta otras cosas que es necesario tener en cuenta. Si escribimos la consabida ecuación

anhídrido carbónico + agua + energía de la luz = materia orgánica + oxígeno,

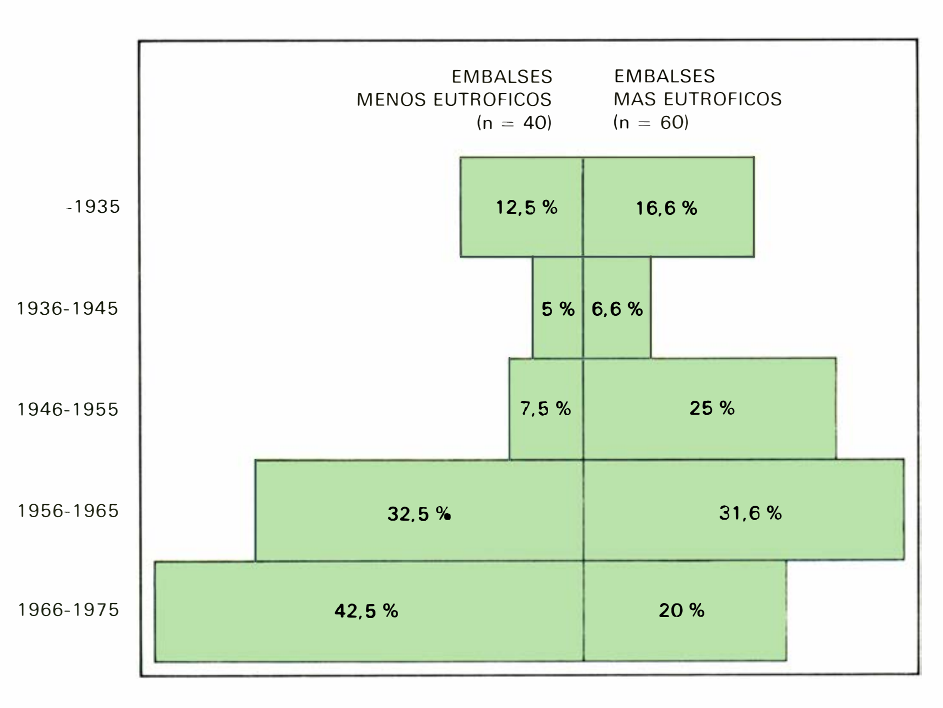
y la leemos de izquierda a derecha y de derecha a izquierda, observaremos que la misma cantidad de oxígeno producida en la fotosíntesis es la necesaria para la oxidación total de la materia formada, incluyendo el que se gasta en el metabolismo de bacterias y de otros organismos. Por lo tanto, si el lago es de gran eficacia biológica, y el agua superficial se sobresatura de oxígeno, huyendo parte de éste a zonas menos concentradas de la atmósfera (por difusión), advertiremos más tarde que se necesitaba ese oxígeno para oxidar la materia orgánica producida y, al no disponer entonces de esa parte, quedará materia orgánica sin descomponer, acumulándose en el agua o incorporándose al sedimento.

La cantidad de materia orgánica que se puede oxidar en profundidad guarda relación con la reserva de oxígeno allí disponible. El oxígeno almacenado en el hipolimnion procede de la estación fría, o de las épocas de transición, cuando hubo mezcla vertical del agua del embalse. Si el hipolimnion tiene, pongamos por caso, 30 m de altura, y retuvo 8 ml de oxígeno por litro durante el período de mezcla vertical, una columna de 1 m² de sección y 30 m de altura habrá retenido unos 240 litros de oxígeno, aproximadamente 350 gramos de oxígeno, que son suficientes para oxidar unos 330 gramos de materia orgánica, equivalentes a unos 130 gramos de carbono, pero no más. Si durante el período de estratificación del lago, e imaginemos que eso ocurre de abril a octubre por ser la época más productiva, la de mayor intensidad luminosa, se transporta del epilimnion al hipolimnion una cantidad mayor de materia orgánica, parte de ésta no podrá oxidarse e irá a posarse en el sedimento. Esta materia orgánica que se incorpora al sedimento confiere a éste un color obscuro característico. Dicha tonalidad sirve para fechar la época de transición de un lago oligotrófico a eutrófico, y así se ha hecho en ciertos lagos suizos a base de muestras sacadas en forma de cilindros de sedimento.

Un lago eutrófico sintetiza en superficie, anualmente, una cantidad de materia viva que contiene de 300 a 600 gramos de carbono por metro cuadrado, de los cuales 200 a 400 gramos por metro cuadrado se producen durante la temporada de mayor producción, que hemos supuesto comprendida entre los meses de abril y octubre, tiempo que corresponde asimismo a la estratificación del embalse. Una parte de esta producción orgánica se recicla en el epilimnion, pero otra parte bastante elevada desciende al hipolimnion. Volviendo al ejemplo anterior, parte de la materia orgánica producida no podrá oxidarse y el lago mostrará los síntomas de eutrofización: déficit de oxígeno y exceso de material orgánico en el hipolimnion.

Los materiales que se acumulan en el sedimento quedan retirados del ciclo, por lo menos durante cierto tiempo. De forma que si no ocurriera nada especial, es decir, si no se diera un nuevo aporte de materiales, al cabo de un año empezaría de nuevo el ciclo con una concentración de nutrientes más baja, tal vez se perderían grados de eutrofización, y este proceso de remisión conduciría con los años a una situación de relativa oligotrofia, en que la producción de materia orgánica es muy baja (hasta 30 a 80 gramos de carbono por metro cuadrado y año), bastando entonces la reserva en oxígeno del hipolimnion para su oxidación total en la misma temporada.

Existen otros mecanismos que coadyuvan a frenar la eutrofización. En un hipolimnion pobre en oxígeno, las bacterias desnitrificantes desarrollan una gran actividad al encargarse de pasar el nitrógeno combinado de los compuestos inorgánicos a nitrógeno en forma molecular, cuya concentración estará en equilibrio con el ya existente disuelto en



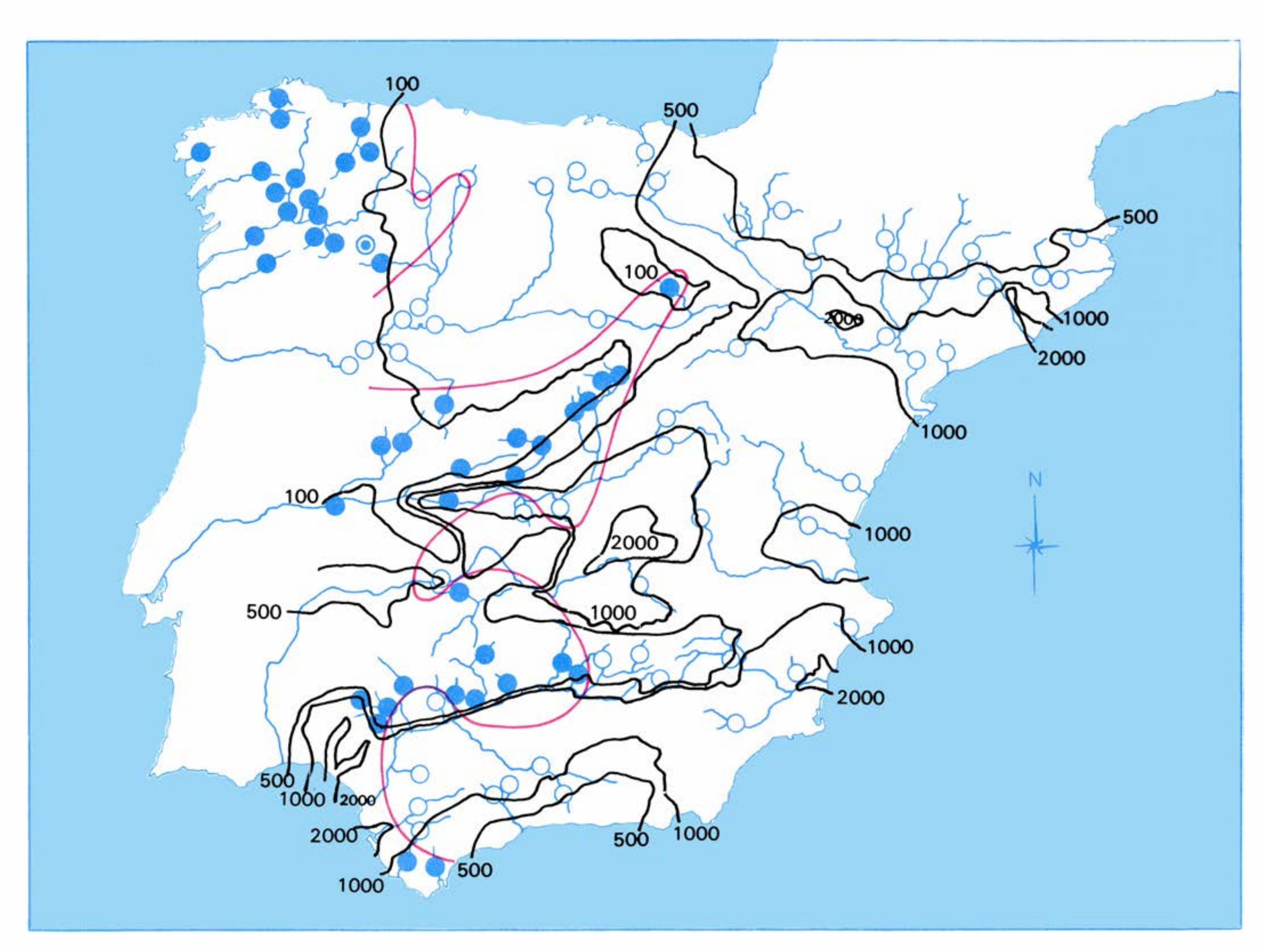
LA EDAD DE LOS EMBALSES es ya un dato que nos permite explicar el grado de eutrofia. Dos grupos de pantanos españoles representativos de los menos eutróficos *(izquierda)* **y los más eutróficos** *(derecha)* **se distribuyen según el año de su construcción** *(izquierda, en ordenadas).*

el agua y con el de la atmósfera. Esta desaparición de los compuestos nitrogenados es notoria en el hipolimnion anóxico (carente de oxígeno) del embalse de Sau *(Véase la ilustración de la pág. 59).* El aumento del pH en el epilimnion, como consecuencia de la asimilación del anhídrido carbónico por las algas en la fotosíntesis, contribuye a que precipiten minerales que contienen fósforo, uno de los nutrientes más importantes. Por tanto, una fracción de este elemento se hurta a futuros ciclos. El descenso en la cantidad disponible de dos nutrientes tan destacados como el nitrógeno y el fósforo, además de la no entrada de nuevos aportes de materiales, inducen a esperar que un lago eutrófico, abandonado a su suerte, se iría convirtiendo en oligotrófico. Puede observarse cómo, en el esquema donde se compara el ciclo de distintos tipos de embalses, los eutróficos se caracterizan porque algunos ciclos parciales se extienden hacia la atmósfera o hacia el sedimento, perdiendo así nutrientes esenciales, lo cual constituye a tales ciclos o bucles en amortiguadores ecológicos, por lo menos en teoría *(véase la ilustración de la pág. 53).*

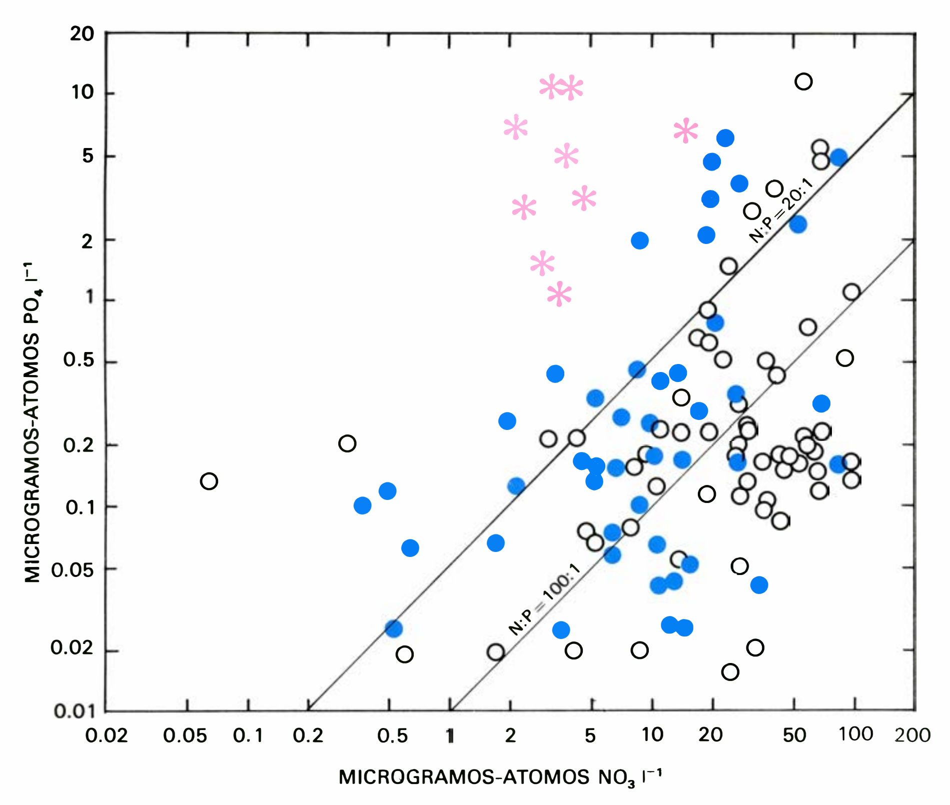
Pero, en realidad, los lagos y embalses eutróficos lo siguen siendo porque no pueden compensar totalmente la eutrofización, ya que continuamente están recibiendo nutrientes, procedentes de la erosión del suelo, los desagües, etcétera, y siguen trabajando a un ritmo intenso. No es raro, especialmente en el caso de los embalses, que reciban directamente materia orgánica a través de un afluente que, por ser de una densidad superior, circula por el fondo. Ello contribuye a agotar la reserva de oxígeno del hipolimnion, con lo cual se reducen las posibilidades de oxidar la materia orgánica producida en el propio embalse; el caso contrario, que un afluente potencie la concentración de oxígeno del hipolimnion, es muy raro. Lo que sí es corriente es que la evacuación preferente de agua del hipolimnion de un embalse disminuya la gravedad de la eutrofización en el propio embalse; sin embargo, ese flujo que sale de la presa, por tener un grado de eutrofización muy elevado, creará problemas en el río al que va a parar.

El mecanismo normal en un lago aislado determinaría que éste pasara de eutrófico a oligotrófico, salvo por el efecto de reducción del volumen del hipolimnion debido a la acumulación de sedimentos que no dependen del lago. Por causa, sobre todo, de la acción humana, el lago o el embalse aparece relacionado con un área más amplia, que influye en el proceso de eutrofización manteniéndolo o acelerándolo. Lo cual es certísimo para la presa que nace bajo el impacto de la civilización, cuya eutrofización es muy rápida.

Ante el dato empírico del tránsito de oligotrofia a eutrofia en las aguas dulces, se creyó que tal sucesión era la natural. Si era natural e inevitable, no habría porqué preocuparse entonces por la eutrofización cultural, por la adición exagerada de nutrientes, que no hacen otra cosa que acelerar la sucesión natural.



DISTRIBUCION DE LOS VALORES DE CONDUCTIVIDAD del agua, expresada en milésimas del recíproco de ohmios, por centímetro; se han distribuido en dos grupos, simbolizados distintamente según que su alcalinidad sea superior o inferior a 1,5 miliequivalentes por litro.



RELACION ENTRE LA CONCENTRACION de nitrógeno y fósforo, expresada aquélla en forma de nitrato y ésta en forma de fosfato, en el agua de los distintos embalses y en distintas fechas. En los embalses más alcalinos, la relación N:P es más elevada, tomando promedios. Se emplean relaciones atómicas. Los símbolos usados para designar los embalses son los que figuran en el mapa que da la localización de los mismos. Los asteriscos denotan algunos embalses meridionales que pudieran estar contaminados por la industria oleícola.

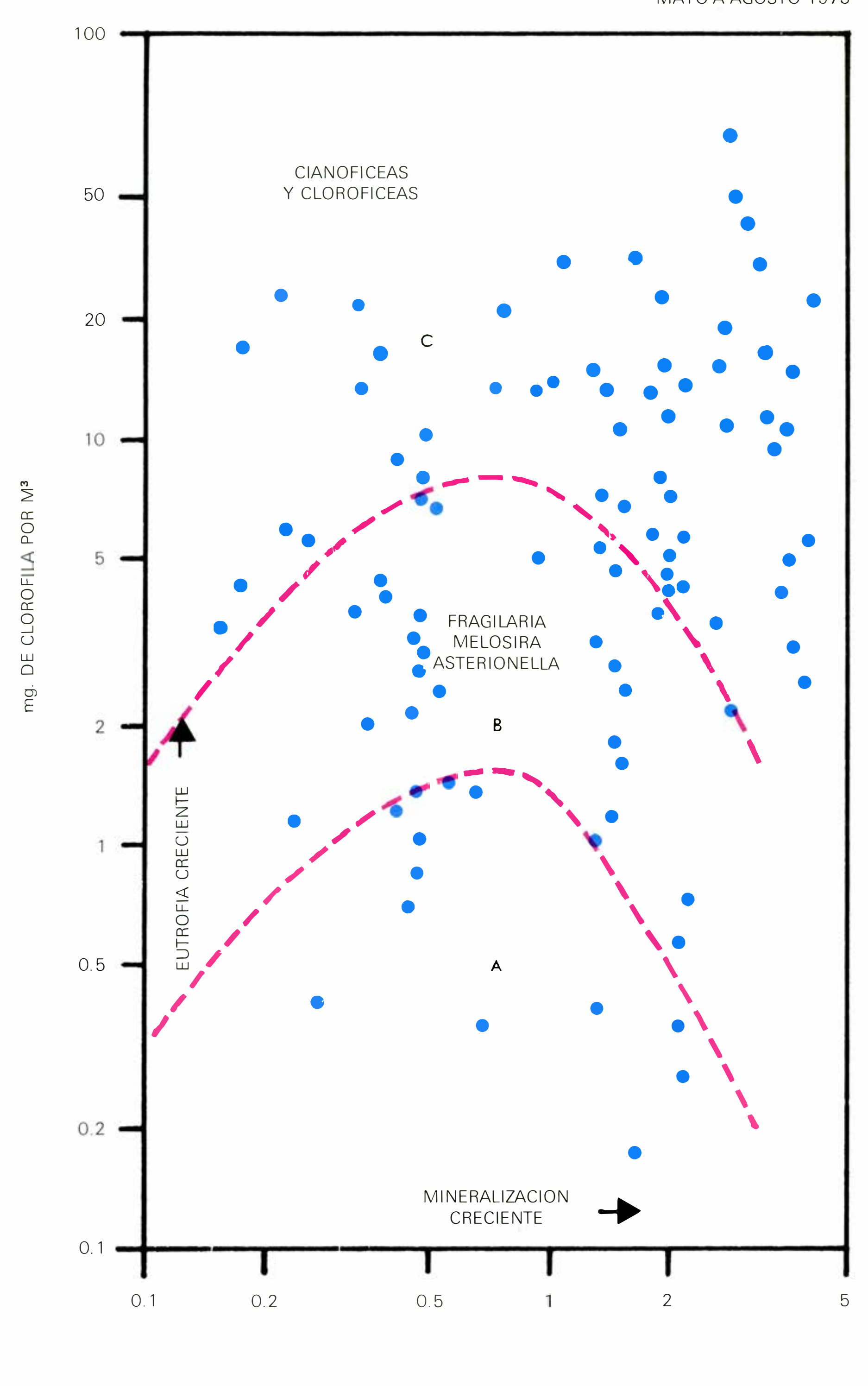
Pero ese proceso no es natural ni inevitable, como prueban las medidas de corrección y mejora del lago Washington, promovidas por Edmondson. De ahí, pues, la constante llamada de atención por parte de los limnólogos contra el vertido de nutrientes, muy especialmente del fósforo, al que se considera como el factor más importante de la eutrofización en lagos y embalses.

Hemos tratado con cierto detalle el proceso de eutrofización, común a lagos y embalses, modificado y particularmente agravado en éstos, no sólo porque es importante en sí mismo, sino también porque ha dado ocasión para presentar algunos principios del funcionamiento de los lagos y embalses de la región templada necesarios para lo que vamos a exponer a continuación. La eutrofización, por supuesto, no agota la problemática de las aguas dulces. Pudiera ser que la tipología de los lagos nos apareciera ahora más confusa. En efecto, teníamos diferencias geográficas o regionales que medíamos según el eje oligotrofia-eutrofia (verbigracia, el contraste entre la montaña escandinava y la llanura báltica), y ahora hemos observado que la tendencia a la eutrofización es común en todos. ¿Hay que pensar, pues, que existen diferencias regionales entre los embalses y lagos que no se deben exclusivamente a la eutrofización? ¿Es, acaso, que la eutrofización presenta características heterogéneas, dependientes de las características peculiares y locales o regionales de los lagos?

Hasta ahora hemos visto la eutrofización por acción cultural. Nos ocuparemos en adelante de las relaciones entre geografía y eutrofia, muy interesantes y raramente tratadas de manera adecuada. La variada constitución geológica de la Península Ibérica y la diversidad en la composición química de sus aguas permiten una respuesta apropiada a los interrogantes formulados más arriba. En lineas generales, España está formada por cuencas terciarias con sedimentos ricos en materiales solubles (incluyendo evaporitas genuinas), limitadas por sierras en las que predominan materiales calizos mesozoicos. Extensas áreas del oeste y, especialmente, del noroeste, así como algunos asomos en los ejes de las principales cordilleras, están formadas por materiales silíceos, paleozoicos o intrusivos (granitos). La climatología muestra una notable correspondencia con la naturaleza geológica, pues la mayor parte de la España silícea recibe 1000 mm o más de lluvia anual, mientras que el resto del país, con rocas y suelos formados por una mayor proporción de materiales solubles, recibe menos agua y repartida de modo que cada año ocurren temporadas de sequía de un mes a seis meses de duración.

Un primer examen de la cuestión nos permitió reconocer la existencia de muchas propiedades de los embalses, manifiestas especialmente en su biología, indiscutiblemente relacionadas con características de las aguas y, por tanto, del substrato geológico. Al propio tiempo, las manifestaciones de la eutrofización variaban de un lugar a otro según dichas características.

Se puede proceder a un análisis generalizado del problema, a partir de la consideración particular de cada embalse entendido como una entidad definida por muchas variables (como lo son también un organismo, un suelo, cualquier ecosistema). Entonces se procede estudiando estadísticamente el comportamiento de las variables, para identificar y reunir las que varían de manera paralela. Tal investigación viene facilitada por los computadores, que posibilitan la cuantificación de los datos, compleja y larguísima. De este tipo de análisis se espera que reduzca el número de variables importantes y simplifique cualquier ordenación según las mismas. Pero las nuevas variables no son sencillas, sino que corresponden a conceptos nuevos que deben describirse en referencia a la combinación habitual de las variables medidas. Los métodos a que me refiero se desarrollaron primero en el campo de la psicología, cuyo éxito se comprende fácilmente pues permitían pasar de datos empíricos (respuesta de individuos de la muestra a un estímulo concreto) a la definición de conceptos nuevos, tales como la agresividad, afectividad u otros rasgos de la personalidad, caracterizados por frecuentes y peculiares combinaciones de las variables originales. Se plantea, pues, el problema de si es posible, siguiendo tal procedimiento, precisar cuantitativamente el concepto de eutrofia y el proceso de eutrofización o cualquier nuevo concepto análogo que pudiera aparecer. Al fin y al cabo, la eutrofia se caracteriza por un conjunto armónico de acciones y reacciones; por consiguiente, las variables que se midan tienen que in asociadas, por lo menos hasta cierto punto. Pero también es cierto que, según qué variable escojamos, influiremos en las constelaciones derivadas de ellas, poniéndose de relieve unos ejes de variación con preferencia a otros. Podremos saber si las variables y los ejes son aceptables o no según la capacidad que presenten para organizar o interpretar otros conocimientos.



GRADO DE MINERALIZACION y grado de eutrofia de los embalses españoles estudiados. La mineralización está representada por la alcalinidad y la eutrofia por la concentración de clorofila en el agua superficial durante el verano. Se han adoptado escalas logarítmicas. Cada punto representa un embalse. Las curvas delimitan tres grupos de embalses de eutrofia creciente (A, B y C). El plancton de los embalses moderadamente eutróficos (A y B) se compone de gran cantidad de diatomeas *(Asterionella, Melosira* y *Fragillaria);* en los más eutróficos (C) predominan las cianofíceas y las clorofíceas). Los embalses de la Península más eutróficos son los de contenido de mineralización media.

En los embalses españoles, las variables medidas son las consideradas típicas en limnología, a saber: temperatura, iluminación, oxígeno disuelto, principales iones (Ca, Mg, Na, K; Cl, CO_3H, SO_4, SiO_4), propiedades que dependen de las concentraciones de los mismos (conductividad, alcalinidad, pH), nutrientes más importantes, representados por los compuestos de fósforo y nitrógeno; a éstos se añaden varios elementos cuya concentración está muy influida por las actividades de los organismos, como el hierro, el manganeso, el silicio, y, por supuesto, los propios organismos, el fitoplancton, medido en número de células, en clorofila y en su capacidad de producir, las bacterias y el zooplancton.

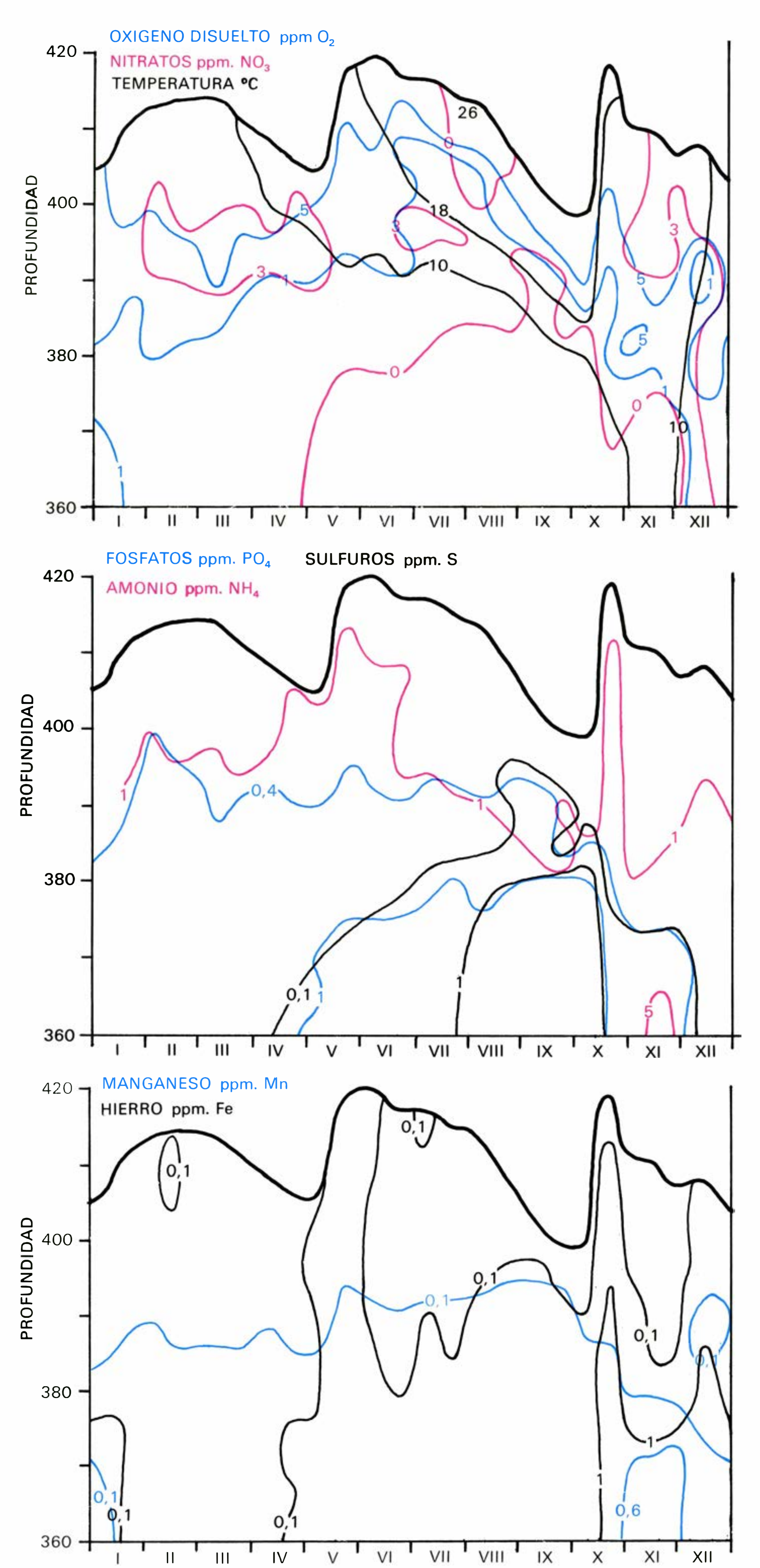
Si nos fijamos en los componentes de proporcionalidad aproximadamente constantes, muy correlacionados unos con otros, forman, en conjunto, una variable más sencilla y general: la mineralización o residuo salino del agua, que se puede medir por la conductividad y, también, de manera apropiada, por la alcalinidad. Esta última es de gran sentido biológico, está correlacionada íntimamente con la mineralización del agua y se define como el exceso de los cationes sobre los aniones fuertes contenidos en el agua. Esas tres variables correlacionadas (mineralización, alcalinidad y conductividad) tienen una marcada distribución geográfica, describiendo a grandes rasgos un gradiente oeste-este, con las aguas de menor residuo mineral –y alcalinidad más baja– en el oeste, y las más mineralizadas –y de alcalinidad más alta– al este. Y así podemos seriar los embalses, o formar grupos discontinuos si se cree más conveniente, desde los que tienen una alcalinidad inferior a 0,3 miliequivalentes por litro, en Galicia, hasta aquellos cuya alcalinidad del agua supera los 3 miliequivalentes por litro, en la España caliza. Las concentraciones de los iones calcio y magnesio, el pH y demás variables típicas varían paralelamente. Ahora bien, cuando se comparan distintos embalses de un entorno geográfico dado se observan desviaciones locales. El pH, además, depende de otros mecanismos amortiguadores. Téngase presente, por último, que la época del año y la profundidad condicionan las correlaciones.

La misma pauta siguen los organismos que habitan los embalses. Así, en el fitoplancton, *Melosira distans* y las diatomeas del género *Tabellaria* se limitan a las aguas del oeste y de los Pirineos, poco mineralizadas, en tanto que las mejor representadas en los embalses más mineralizados del este serán *Ceratium hirundi-*

nella y pequeñas *Cyclotella.* Análogamente, hallaremos diferencias comparables en la distribución zonal del zooplancton: *Ploesoma hudsoni, Conochilus* y *Thermocyclops dybowskii* se limitan al país silíceo, principalmente del oeste de la Península, mientras que las *Synchaeta* del grupo *triophthalma* y *Daphnia magna* son características de las aguas más mineralizadas.

Existe una cartografía oficial de la calidad de las aguas españolas, independiente de los datos obtenidos en los mismos embalses, basada en medidas de conductividad, alcalinidad y otras variables, obtenidas por los servicios estatales de control de la calidad del agua de los ríos. En el mapa de la página 56 se superponen estos datos, referidos a los ríos, con una repartición de los embalses en dos grupos. Un mapa de esta índole tiene interés práctico, pues las características de los procesos de eutrofización varían según las regiones y las medidas aconsejables para la conservación de las aguas no tienen por qué ser coincidentes.

Podemos ahora precisar con mayor exactitud la naturaleza peculiar de la eutrofización en los embalses, y sus diferencias de la eutrofización que advertíamos en los lagos. El flujo intenso a que suelen estar sometidos y la posición del efluente, que según los casos puede evacuar agua del hipolimnion, modifican los términos en la formulación del problema de la fertilización de sus aguas. Por supuesto, la aportación de nutrientes, complementados frecuentemente por materia orgánica, conduce a una elevada producción de plancton y agotamiento ulterior del oxígeno. La secuencia de los cambios en el embalse de Sau, que se ilustra en la página 59 podría ser típica de cualquier lago eutrófico, a pesar de las fuertes variaciones en el nivel del embalse: el agua se estratifica, se agota el oxígeno, se desnitrifica el hipolimnion, aparecen sulfuros y luego amoníaco y se libera fósforo, manganeso después y, por último, hierro del sedimento. A la liberación del fósforo contribuye la presencia de sulfuros; el manganeso se desprende cuando el potencial de reducción y de oxidación del agua alcanza un valor determinado; y el retardo con que apa-



VARIACION DE ALGUNAS PROPIEDADES del agua en el embalse de Sau, provincia de Barcelona, en función del tiempo (meses del año 1970; *de izquierda a derecha*) y de la profundidad (hacia arriba y expresada como niveles fijos sobre el del mar), de manera que las fluctuaciones del nivel del embalse se reflejan con el contorno del límite superior del presente gráfico (según Antonio Vidal).

rece la concentración máxima del hierro presente en el agua se debe a la asociación de parte del mismo con formas de materia orgánica.

Esos fenómenos ilustrados consisten esencialmente en intercambios entre el agua y el sedimento profundo. Su enorme interés no radica sólo en que retiran parte del fósforo que entra en el lago o embalse, en una fracción no linealmente proporcional a la carga, sino también en el fenómeno inverso, en regular el retorno del fósforo al lago, retorno que es capaz de prolongar el estado de eutrofia por un largo período de tiempo, incluso después de haber tomado medidas correctoras. Para muchos elementos existe además la posibilidad de irse acumulando lentamente en el sedimento del embalse. El embalse de Sau, según análisis de Vidal, acumula más de diez toneladas de manganeso cada año en su sedimento. La liberación brusca de gran cantidad de ese elemento, y de otros, bajo circunstancias excepcionales, es un riesgo que no debe olvidarse.

No existen en la Península embalses equivalentes a lagos oligotróficos, pero los hay con distintos grado de eutrofización. Por razones de orden práctico, podemos agruparlos en menos y en más eutróficos, tomando como límite o referencia para tal división los valores siguientes, considerados promedios: 15 mg de fósforo por m³, en lo que respecta a la concentración de fosfato; 140 mg de nitrógeno por m³ de concentración de nitrato; 5000 células por ml, o 5 mg de clorofila por m³, en lo que respecta a la concentración de fitoplancton; 3 mg por m³de agua, y hora, respecto a la producción primaria o asimilación de carbono por el fitoplancton. Si los valores observados rebasan repetidamente a los indicados, se trata de un embalse muy eutrófico, que planteará problemas para ciertos usos de su caudal. La profundidad a la que desaparece la visión de un disco blanco sumergido en el agua (disco de Secchi) proporciona una idea inmediata y aproximada del grado de eutrofia, sabiendo que si desaparece a una profundidad inferior a los tres metros la intensidad de la eutrofia es muy elevada.

Cuando un embalse se llena por primera vez, la fertilidad de su agua aumenta rápidamente debido al suelo y a la vegetación que haya podido quedar sumergida. Viene luego una detención o un retorno a condiciones menos eutróficas. Ahora bien, con la aportación de los ríos se origina una eutrofización progresiva irreversible. Como pone de manifiesto la distribución de los embalses estudiados en la figura de la página 55, los pantanos tienden a ser más eutróficos cuanto más antiguos son; por tanto, si el conjunto de los embalses no parece encontrarse en una situación muy grave en lo que a este aspecto se refiere, se debe a que son de construcción relativamente reciente.

La ordenación de los embalses por su grado de eutrofia no aparece tan sencilla como su ordenación por el grado de mineralización de sus aguas. Si consideramos, a la vez, la fertilidad, que se aprecia por la concentración de clorofila en superficie, y la mineralización del agua de los embalses *(véase la ilustración de la página 58)*, se deduce que los embalses más eutróficos son preferentemente, los de mineralización intermedia. Ello no deja de ser paradójico; se comprende bien que las aguas con pocas sales disueltas sean poco fértiles, pero no resulta tan inmediato que las aguas más mineralizadas no sean, a la par, las más productivas. Aunque la solución no es simple, puede encontrarse una parte importante de la misma en la tendencia de los fosfatos a formar minerales insolubles en las aguas alcalinas. Por consiguiente, una fracción de este nutriente se hurta a su posible utilización en el desarrollo del plancton. Como ocurría en los lagos, las características de la eutrofización dependen no sólo de la aportación de nutrientes, sino también del resto de la composición del agua.

Si existe ese secuestro del fósforo y no lo hay del nitrógeno, la relación entre nitrato y fosfato se halla relativamente aumentada en los embalses de agua más alcalina *(véase la ilustración de la pág. 57)*. En esas condiciones se desarrolla un plancton normal, es decir, que dispone del nitrógeno presente en el agua formando parte de compuestos inorgánicos y lo utiliza así. En los embalses del oeste y de las montañas, de aguas con menos residuo mineral, el fosfato aportado queda en solución; si la presa es muy productiva y se agota el oxígeno en el hipolimnion, las bacterias se encargan de la desnitrificación, pasando nitrógeno a forma gaseosa o molecular. En estas condiciones, si queda fósforo disponible o se aportan nuevas cantidades de este elemento, cobran ventaja los organismos del plancton que son capaces de fijar nitrógeno gaseoso, como las cianofíceas, en especial las provistas de heterocistes, que son células diferenciadas en las que se localiza la función fijadora de nitrógeno. Las algas cianofíceas son tóxicas algunas de ellas y casi todas contribuyen a crear problemas o dar mal sabor al agua, si llegan hasta las estaciones depuradoras.

La proliferación de las cianofíceas es un ejemplo de las complejidades de la limnología y muestra en la práctica que no pueden aplicarse los mismos principios de vigilancia y tratamiento de aguas en todas partes. En un país silíceo hay que evitar la llegada de fósforo a los embalses y preocuparse menos del nitrógeno, cuya limitación resulta contraproducente, pues la escasez de nitrógeno combinado estimula el desarrollo de las cianofíceas, con los problemas resultantes. En las aguas más mineralizadas, no habrá que preocuparse tanto por separar el fósforo, a través de lo que se llama un tratamiento terciario en la depuración, dado que buena proporción del fosfato precipita con facilidad en el embalse o en los mismos afluentes, antes de llegar al mismo; sí habrá que preocuparse ante la posible liberación del fósforo (y de metales) del sedimento en que se han ido acumulando; ese desprendimiento acontece o se facilita cuando aparecen sulfuros en el medio. Lo cual se presenta reiteradamente en las aguas más mineralizadas y ricas en sulfatos, cuando se aporta materia orgánica, utilizable como fuente de energía por las bacterias que reducen el sulfato a sulfuro. En conclusión, es obvio que la importante función del fosfato y de la materia orgánica depende de las características del embalse. El mecanismo siempre es fundamentalmente el mismo, pero la presencia en mayor cantidad de otros iones –principalmente de calcio– desvía el acontecer según secuencias diferentes.

En el estudio de los embalses españoles ya se han censado, sólo en el plancton, más de 700 especies de cianofíceas y otras algas, 113 de rotíferos y 63 de crustáceos; en el bentos, se han reconocido larvas de 72 especies de una sola familia de dípteros, los quironómidos. Esto da idea de la variedad de formas de vida y de la complejidad de los factores que han de definir la presencia y el predominio de unas u otras. Muchas especies se encuentran bien distribuidas en diversas aguas de la Península y han tenido fácil acceso a los embalses. Un gran número de organismos de agua dulce resisten, en alguna forma, la desecación, y ellos o sus huevos o gérmenes pueden ser transportados con el polvo atmosférico o las aves acuáticas. Pero, aunque no faltan explicaciones *a posteriori*, la colonización de los embalses españoles por sus pobladores actuales tiene que ser una historia llena de motivos de interés.

Los embalses españoles, en cuanto medio ambiente, difieren bastante de las aguas naturales del país, y se asemejan más a los lagos del margen de los Alpes y a los embalses situados más allá de los



COLONIZACION de los distintos embalses españoles por organismos animales que se encuentran en el plancton de los mismos y que, verosímilmente, son de origen extraño a la Península, predominando, según la ilustración, los embalses situados en el NO-SO (según Juan Armengol).

Pirineos. No es de extrañar que la distribución de los organismos refleje estos hechos de algún modo. Las tres diatomeas más abundantes en los embalses de eutrofización mediana, *Asterionella formosa, Fragilaria crotonensis* y *Melosira granulata,* son muy raras en las aguas naturales de la Península, pero abundan en los lagos europeos, y uno piensa espontáneamente si cepas que se desarrollan en nuestros embalses no vendrán de más al norte.

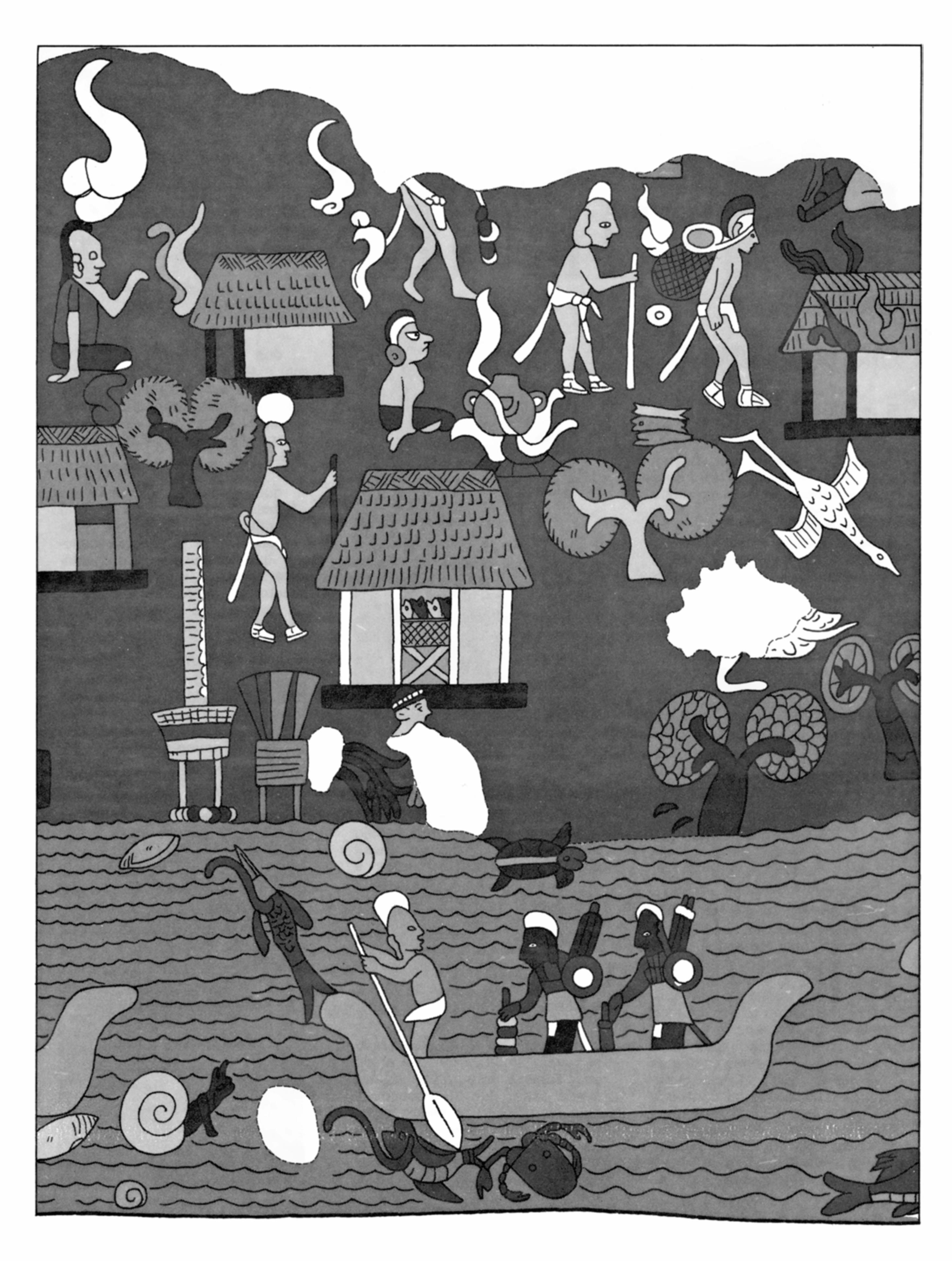
Esta impresión de colonización desde fuera tiene mayores pruebas de verosimilitud en relación con los animales, cuyas especies suelen distribuirse en áreas más restringidas y mejor delimitadas que las algas. En embalses de España se encuentra la medusa de agua dulce, *Craspedacusta sowerbyi* que, considerada originaria de China, aparece hoy más o menos dispersa por todo el mundo, casi siempre en embalses. Hay varios animales del plancton que faltan en las aguas naturales españolas, cuya presencia en los embalses puede atribuirse a una colonización reciente *(véase la ilustración de la pág. 61).* Esto es indudable para una *Daphnia* o pulga de agua americana, y posiblemente también sea cierto para determinadas formas europeas del mismo género, caracterizadas por un desarrollo especial de la cabeza en algunas generaciones nacidas durante el verano. Tales *Daphnia* no se conocían al sur de los Pirineos. No es sorprendente la presencia de formas tropicales, como *Ceriodaphnia cornuta* y *Keratella tropica* en nuestros embalses. *Copidiaptomus steueri* se halla actualmente en muchos embalses del oeste de la Península; fuera de ella, es específica de los países del Mediterráneo oriental. Otro diaptómido, *Diaptomus castaneti,* es, sin embargo, un endemismo descrito en el lago de Sanabria, que ha colonizado luego una serie de embalses de la región.

La constitución de las comunidades de organismos que pueblan los embalses sigue siendo un proceso activo. En el corto tiempo de su estudio se han puesto de manifiesto sustituciones de unas especies por otras, que descubren interesantes fenómenos de acceso y competencia.

La fauna del fondo de los embalses presenta problemas de interés igual o superior. Sus representantes más importantes son los gusanos oligoquetos y las larvas de quironómidos. Los oligoquetos se sustentan, en buena parte, de los detritus procedentes de la periferia de los embalses, en tanto que los quironómidos se nutren en una mayor proporción, del fitoplancton que se sedimenta. De manera que la importancia relativa de los distintos grupos pone de manifiesto las condiciones de alimentación forzada del embalse. Pero la presencia y distribución de unos u otras especies refleja, sutil y precisamente, las características del embalse, y no sólo instantáneas, sino integradas sobre cierto tiempo.

La composición del fitoplancton se puede utilizar asimismo para distribuir los embalses en grupos, que tienen validez y sentido práctico en relación con su manejo y gestión. Los pioneros de la limnología utilizaron la composición de las poblaciones bentónicas de quironómidos para la distinción de tipos de lagos: en el fondo oxigenado de lagos oligotróficos, o a menos profundidad en los embalses, se encuentran larvas verdosas o amarillentas; las larvas de los *Chironomus,* de intenso color rojo por la presencia de hemoglobina, son propias de los fondos con poco oxígeno de los lagos eutróficos. La fauna del fondo falta sólo cuando la concentración de sulfhídrico es considerable.

En resumen, el estudio ecológico de los embalses representa la integración de numerosas observaciones en campos muy diversos de la ciencia. El punto de vista que las enlaza es esencialmente dinámico, y conduce frecuentemente a formulaciones y expresiones cuantitativas, más apropiadas para el desarrollo de la ciencia, es decir, las más apropiadas para ser comprobadas, y declaradas luego ciertas o falsas. Tanto las necesidades crecientes de agua, como otros problemas ecológicos relacionados con la ocupación de los valles por embalses, conducen a una inmediata aplicación de los conocimientos científicos, hasta el punto que nos movemos en un campo en que es difícil deslindar dónde acaba la ciencia pura y empieza la práctica. Continuamente se construyen nuevas presas y aparecen en las viejas nuevos problemas; y en todas ellas se desarrolla la vigorosa actividad de la vida, colonizando, ocupando, compitiendo, en todos los nuevos ambientes en un proceso incesante de sucesión y evolución. Son ilimitadas las perspectivas que, desde el punto de vista científico, ofrece el estudio de los embalses, y hay que señalar expresamente el interés práctico que tiene conocer el intercambio de elementos químicos entre el agua y el sedimento junto a los problemas generales de la eutrofización. Allí está la clave de muchos mecanismos reguladores del desarrollo de la vida en los embalses y, en el caso particular de proximidad local e interacción entre embalses e instalaciones de energía nuclear, pueden incidir directamente en la salud.



CANOA EN MISION EXPLORADORA, con un solo remero y dos guerreros armados, bordeando la costa; el grabado pertenece a una pared pintada del Templo de los Guerreros, en Chichén Itzá. Los guerreros se representan portando armas de origen mexicano: dardos en la mano y un carcaj cerca del escudo. Probablemente representan la vanguardia de los toltecas, quienes reconstruyeron Chichén Itzá durante el siglo X.

El surgimiento de una clase mercantil maya

Se creía que los siglos inmediatamente anteriores a la conquista española conformaban un período de recesión y decadencia en la civilización maya. La revisión de las fuentes desvela, por el contrario, que había nuevos líderes que perseguían nuevas metas

Jeremy A. Sabloff y William L. Rathje

El siglo VIII fue un período turbulento en la historia mundial. En el Viejo Mundo, la victoria de Carlomagno en Tours detuvo el avance musulmán en Europa, pero los musulmanes, después de sitiar Constantinopla por segunda vez, todavía tuvieron la fuerza suficiente para seguir considerando el Mediterráneo como su mar privado, aplastar la dinastía Maitraka en la India y hacer retroceder la frontera de la dinastía Tang más allá del Turquestán chino.

Aunque nadie del Viejo Mundo estuviese enterado de ello, en el Nuevo, concretamente en lo que son hoy México y Guatemala, lo que fuera un período de prosperidad económica y de florecimiento cultural estaba terminando en un desastre. El gran centro urbano de Teotihuacán, situado en las altiplanicies de México, fue saqueado e incendiado alrededor del año 700 d. C.; en la parte meridional de las llanuras de Yucatán, la civilización clásica maya se estaba hundiendo, sus magníficos centros rituales dejaron de funcionar y la región se despobló. La situación que siguió a este desastre en el Nuevo Mundo es el tema que nos ocupa.

Muchos estudiosos de la prehistoria maya caracterizan el período que transcurre entre el colapso de los llanos meridionales y la llegada de los conquistadores españoles como una era de lenta decadencia en toda el área maya, tanto en los llanos meridionales como en los septentrionales. Nos han llegado muy pocos escritos del Nuevo Mundo anteriores a la entrada de los españoles, en el siglo XVI y lo que nos ha llegado sólo es inteligible en parte. Por tanto, los que emprenden la tarea de reconstruir la historia precolombina dependen en gran medida de la prueba arqueológica.

Tal tipo de pruebas admite múltiples interpretaciones. Por ejemplo, veinte años atrás, Tatiana Proskouriakoff presentó una comunicación sobre su trabajo de varias campañas en Mayapán, en el Yucatán septentrional, financiado por la Carnegie Institution, de Washington. Una gran ciudad amurallada, Mayapán, floreció desde mediados del siglo XIII hasta mediados del XV. La denominación más neutra que los científicos dan a este período de Mesoamérica es la de *postclásico tardío;* pero hay muchos que lo llaman simplemente *el período de decadencia.* Los investigadores financiados por la Carnegie hallaron que la mayoría de las muestras de civilización avanzada de los mayas durante el período clásico temprano y el tardío (300-800 d. C.) no existían en Mayapán prácticamente. Las imponentes estructuras en mampostería de esmerada construcción y la escultura monumental en piedra, adornada con inscripciones, brillaban por su ausencia; ni siquiera una artesanía comparativamente modesta, como la producción de cerámica policromada. Por ello, Proskouriakoff tituló su comunicación "La muerte de una civilización".

Las investigaciones arqueológicas realizadas recientemente en la isla de Cozumel, frente a la costa oriental de Yucatán, nos llevan a una interpretación completamente distinta de este período de la historia maya. En lugar de decadencia y estancamiento cultural, vemos una reorientación cultural: una transferencia de la autoridad a otras manos y, por consiguiente, la búsqueda de nuevos objetivos. Antes de presentar los datos obtenidos en Cozumel y de ofrecer una nueva interpretación de los anteriores, será de utilidad que describamos el panorama general.

La principal ciudad-estado de América del Norte surgió en el alto valle de México, a principios de la Era Cristiana. El lugar de sus ruinas se conoce todavía por Teotihuacán. Del 250 al 600 d. C., los populosos barrios de la ciudad, unos dedicados a los servicios religiosos, otros a los burocráticos y muchos a los mercantiles e industriales, alcanzaron su máximo desarrollo. De acuerdo con los cómputos estadísticos de domicilios realizados por René Millon, de la Universidad de Rochester en la década de los 60, habitaban durante su época de máximo esplendor no menos de 75.000 personas; el número real puede haber sido 125.000. Entre la población activa había artesanos que producían para la exportación manufacturas tan valiosas como la cerámica, objetos de obsidiana y bisutería diversa de piedras semipreciosas. Los artesanos constituían un elemento clave en una economía que exportaba sus productos a todo el territorio, desde Teotihuacán hasta puntos de México y Guatemala situados a cientos de kilómetros. Si las exigencias del comercio eran entonces las mismas que se dieron más tarde, el movimiento comercial hubo de surtir a la ciudad de abundancia de productos, exóticos unos (plumas, resina aromática y cacao) y otros corrientes (sal y tejidos de algodón).

Durante el período en que Teotihuacán disfrutó de los beneficios de un comercio exterior de amplia irradiación, en las llanuras de Guatemala y del sureste de México tenía lugar una expansión parecida *(véase ilustración en la pág. 71).* El pueblo responsable de este floreci-



LA ECONOMIA estuvo dominada, desde los comienzos de la Era Cristiana hasta la conquista española en el siglo XVI, primero por los mercaderes de Teotihuacán y luego por los comerciantes de los alrededores de Tollán. Ambas ciudades estaban situadas en los altos mexicanos. Las rutas comerciales eran terrestes. Al caer Teotihuacán, un grupo maya, los putún, inauguró el comercio por vía marítima cuya ruta se muestra en la ilustración de la página siguiente.

miento meridional era el maya. Su historia se remonta, por lo menos, hasta mediados del primer milenio a. C., pero hasta el período temprano clásico (300-600 d. C.) no se empezaron a grabar, literalmente, en piedra los testimonios de sus grandes hazañas. En Tikal, la más sobresaliente de las ciudades maya clásicas de los llanos meridionales, existe un monumento de piedra fechado con una cifra equivalente al año 292 d. C.

La costumbre maya de fechar los monumentos que se erigían ha supuesto una gran ventaja para los investigadores. Ha permitido la reconstrucción de una cronología bastante próxima a las actividades de los mayas. Por ejemplo, las fechas de los monumentos en los llanos meridionales indican que dicha población trabajó allí por la misma época en que los mayas de los llanos septentrionales construían centros de la importancia de Dzibilchaltún. Las fechas de los monumentos nos informan también de que el siglo VII, centuria considerada tradicionalmente como de transición del clásico temprano al clásico tardío, fue una época de cambio que reflejó algo más que los acontecimientos locales de los mayas.

En México hubo desórdenes durante casi todo el siglo VII. Precisamente, a finales del mismo, Teotihuacán fue incendiada, y el efecto de su destrucción se dejó sentir ampliamente. Mientras se desvanecía la influencia de Teotihuacán sobre Mesoamérica, la civilización maya alcanzó todavía nuevas cimas durante el período clásico tardío (600-800 d. C.). Los centros ceremoniales antiguos aumentaron su capacidad de un modo significativo y se fundaron otros nuevos; se levantaron construcciones típicas de la civilización clásica, como canchas para jugar a la pelota, palacios y grandes templos. Se alcanzó un nuevo dominio de la técnica y del estilo, tanto en la escultura monumental como en el dibujo de la cerámica policroma. La población aumentó rápidamente en los grandes centros mayas. Por ejemplo, se estima que, al principio de ese período, la población de Tikal oscilaba entre las 30.000 y las 50.000 almas.

Tales líneas de fuerza condujeron a la civilización maya a su punto más alto, pero establecieron asimismo los cimientos de su repentino colapso. En poco más de un siglo, entre 770-890 d. C., se abandonaron los centros ceremoniales; no hay un solo monumento en los llanos meridionales que lleve inscrita una fecha posterior al año 900 d. C. Las causas del derrumbamiento de la civilización clásica en los llanos meridionales y del despoblamiento del área son bastante complejas y no nos conciernen aquí. Parece tratarse de un incremento abusivo de los impuestos sobre la zona de los llanos y presiones externas de tipo económico y militar.

Sin embargo, contemporáneamente al derrumbamiento de la civilización maya, otra ciudad-estado de los altiplanos de México comenzaba a llenar el vacío dejado por la destrucción de Teotihuacán. Los investigadores mexicanos identifican dicha ciudad con Tollan (Tula moderna), capital de los toltecas, cuya fundación data tradicionalmente del año 900 d. C. Menor, y de inferior importancia en el terreno económico, que Teotihuacán, Tollan ejerció no obstante una gran influencia por su poderío militar.

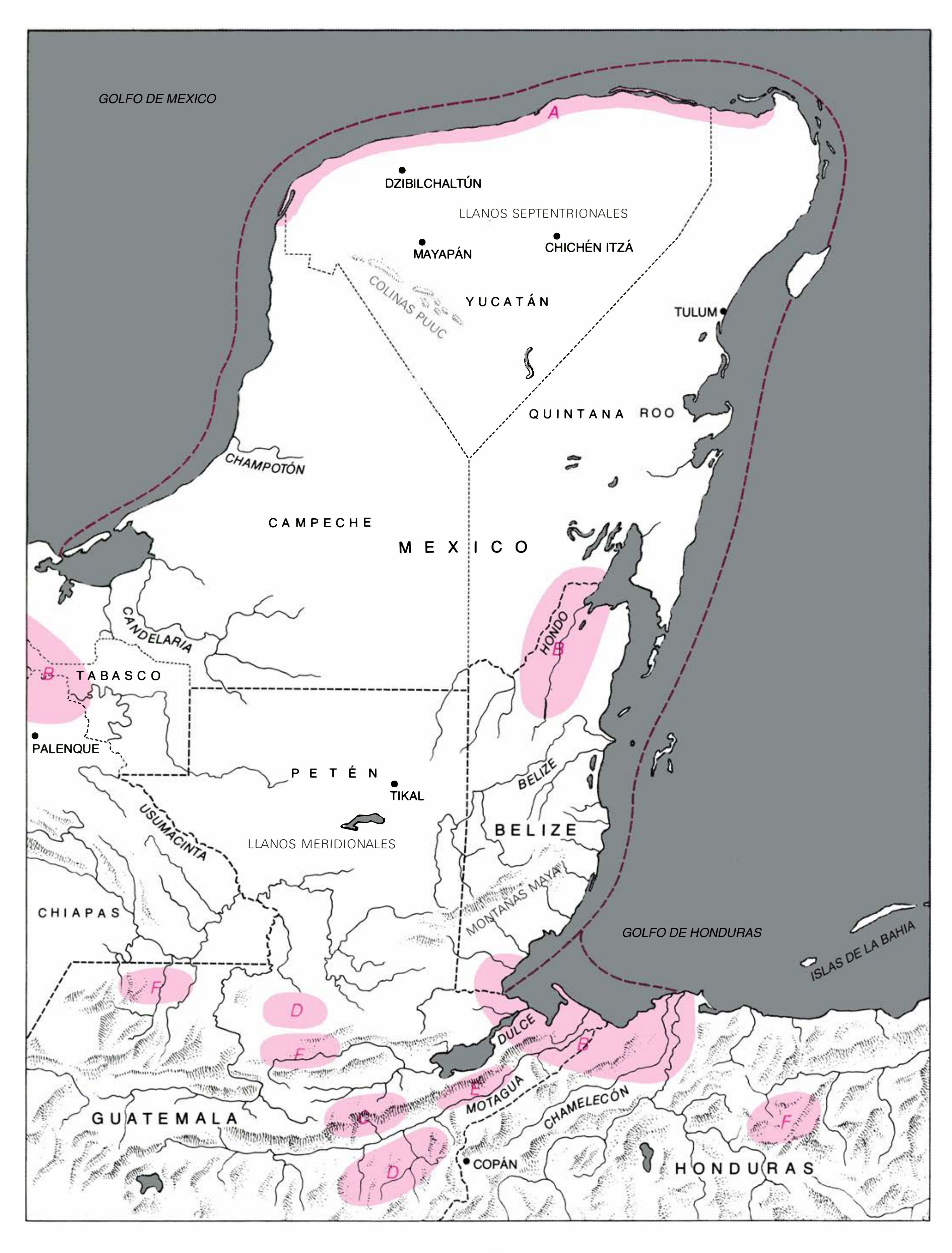
Tollan extendía su comercio de importación hacia el oeste, el norte y el sur del valle de México. Respecto al comercio con el sur, los análisis arqueológicos de la cerámica importada, excavada en Tollan, indican que las rutas mercantiles que cubrían el país se extendían hasta Costa Rica y Nicaragua. Y en lo que toca a sus más próximos vecinos, los mayas, la prueba más espectacular de la influencia tolteca se encuentra en Chichén Itzá.

Construida originariamente por los mayas del clásico tardío, Chichén Itzá se amplió y reconstruyó en el estilo tolteca durante el siglo X. Ese gran centro representa, a buen seguro, el peso de la influencia tolteca en la arquitectura maya. Sin embargo, el influjo de los activos emigrantes de Tollan no se limitó a la arquitectura. Al filo del primer milenio de nuestra era, los toltecas habían asumido el control político, casi total, de los llanos septentrionales. La interpretación tradicional de las fuentes arqueológicas mayas acepta esta fecha como punto de arranque del período de decadencia denominado postclásico temprano.

Unos doscientos años después, Tollan fue incendiada, como lo fuera Teotihuacán quinientos años antes. Hacia 1224 Chichén Itzá fue abandonada, y el poder tolteca se extinguió en Yucatán. Para la interpretación tradicional, el declive maya del postclásico temprano se acentuó por estas fechas, abocando en el desplome maya del postclásico tardío.

El escenario está ya preparado. Los actores que entran son los mayas de los llanos costeros del golfo, de los estados actuales de Tabasco y Campeche. Los análisis de documentos posteriores a la conquista, realizados por investigadores como Ralph L. Roys, France V. Scholes y J. Eric S. Thompson, identifican este pueblo como el grupo conocido por los putún. Hablan un dialecto distinto del maya yucateca, que se escucha por casi todo Yucatán; es el maya chontal, uno de los dialectos del grupo maya cholan. Los putún ocuparon un nicho social nuevo, a saber, el de mercader-navegante.

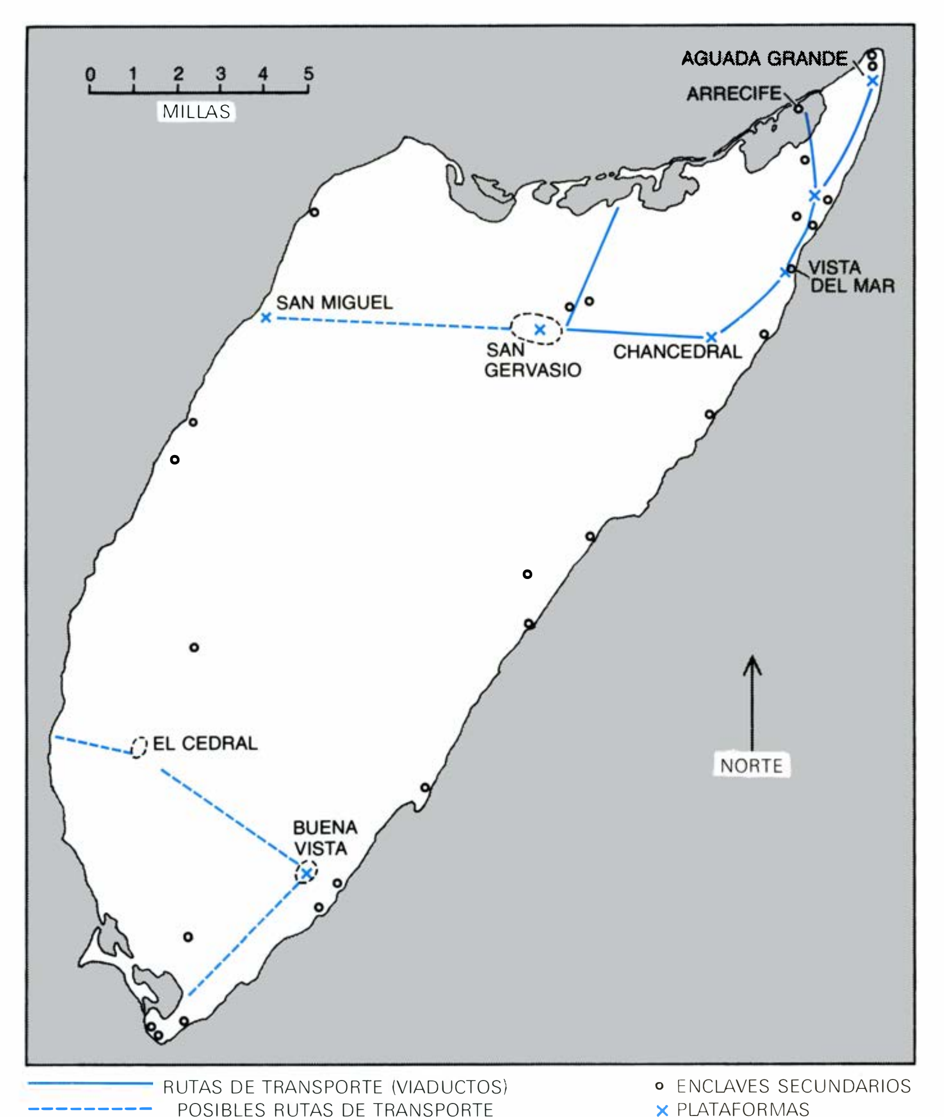
No es fácil sobrevalorar la importancia del comercio entre los pueblos de Mesoamérica. La influencia económica de Teotihuacán llevó a una creciente secularización de la vida y al desarrollo de una economía de mercado en todas partes, no sólo en los altos de México. Al



RUTA MARITIMA PUTUN

A SAL
B CACAO
C JADE
D OBSIDIANA
E COBRE
F PLUMAS DE QUETZAL

LA RUTA MARITIMA PUTUN unía la costa que bordea Yucatán, próxima a las rutas comerciales terrestres del altiplano de México, con puertos del golfo de Honduras. Cozumel, isla situada frente a la costa oriental, servía de cabeza de puente, muy importante para los mercaderes que comerciaban con sal, telas, cacao, jadeíta, plumas, obsidiana y cobre.



ISLA DE COZUMEL, con más de 450 km^2. de superficie. Tiene 34 yacimientos precolombinos. Los enclaves de la costa alertaban probablemente a los habitantes de la isla sobre la proximidad de navíos. Las plataformas de almacenamiento, cuya altura superficial les permitía permanecer secas durante las inundaciones estacionales, comunicaban con puntos de embarque costeros mediante un sistema de viaductos válido para cualquier período del año. El santuario de la diosa Ix Chel, que atrajo tantos peregrinos mayas hasta Cozumel, pudo hallarse situado en San Gervasio.

mismo tiempo, el impulso inicial dado al comercio de bienes de lujo se había ido transformando en una demanda de bienes de uso diario que podían obtenerse de lejanas procedencias. El comercio, sin embargo, había sido casi exclusivamente terrestre. En una economía que carecía de vehículos de ruedas y de animales de carga, ello significaba que la capacidad de transporte estaba limitada a lo que pudieran resistir las espaldas del comerciante y su cuadrilla, contratada o esclava.

Con la caída del imperio económico y político de Teotihuacán, a principios del siglo VIII, los mercaderes-navegantes putún se encontraron en una posición única. No había quedado ninguna organización para hacerse cargo de la explotación y exportación de la totalidad de las materias primas y otros productos de la gran región maya. Por ejemplo, la sal constituía un producto de exportación importante de los mayas. Se extraía de los panes naturales de sal existentes a lo largo de la costa del golfo de la península de Yucatán y en las lagunas de la costa septentrional. Los granos de cacao eran otro ejemplo. La bebida que se hacía de ellos fue el regalo que los mexicanos hicieron al Viejo Mundo; el nombre de "chocolate" es una corrupción de la palabra nahuatl para la bebida. Los granos tenían otra finalidad: se usaban como moneda por toda Mesoamérica. Los primeros exploradores españoles pudieron comprobar cómo el coste de un esclavo vigoroso era de 100 granos de cacao. Tabasco, en el límite de los llanos meridionales, producía abundante cacao. Un tercer artículo de comercio lo constituía el tejido de algodón; los tejedores mayas sobresalieron en la manufactura de los tejidos de algodón. Y un cuarto producto era la miel, que constituia una de las grandes exportaciones mayas.

Los putún se encontraban en el lugar y en el tiempo precisos para continuar el comercio interrumpido de estos y de otros productos; su oportunidad no terminaba allí; al ser navegantes podían transportar las mercancías con mayor eficacia que cualquier otro comerciante de tierra firme. No se trataba sólo de un caso de mejores medios de transporte: los putún supieron planear con oportunidad y eficacia. Los alfareros putún, por ejemplo, fueron pioneros de la producción en serie de formas simples. Se especializaron en una cerámica dura, de paredes delgadas, que los arqueólogos llaman Naranja Fino. La cerámica no se producía en serie, sin ulteriores consideraciones; frecuentemente le daban una forma especial para que unas piezas pudieran introducirse dentro de las otras y así facilitar su embarque *(véase la ilustración de la pág. 74)*.

La ruta comercial de los putún era de cabotaje y de larga irradiación y llegó a extenderse de Tabasco al golfo de Honduras, con la isla de Cozumel como cabeza de puente. Se sabe poco de las embarcaciones de los putún, descritas por los primeros visitantes europeos como troncos vaciados, de medidas variables. Los relatos españoles hablan de canoas que podían transportar hasta cincuenta hombres. Las representaciones de las canoas mayas, realizadas por artistas propios, figuran siempre embarcaciones más pequeñas. Esto puede deberse a que los artistas dibujaban las canoas con la capacidad necesaria para acomodar a sus ocupantes, siendo siete el número más alto de pasajeros que se representan. En cualquier caso, por el tiempo en que arribaron los primeros españoles, el dialecto putún se había convertido en la lengua franca del comercio en toda el área maya. La reputación viajera de los mercaderes era tal que Thompson les llama "fenicios del Nuevo Mundo". La literatura de la exploración europea narra un encuentro entre un barco español y una canoa de las que navegaban por el óceano, y que podía muy bien pertenecer a un comerciante putún. El contacto ocurrió durante el cuarto viaje de Colón, que estaba enfilando una de las islas de la Bahía, frente a la costa de Honduras. A últimos del mes de julio de 1502, su barco llegó a la altura de una gran canoa, de unos dos metros y medio de ancho y "la longitud de una galera", impulsada por

unos veinte remeros. El capitán, probablemente el comerciante-propietario, era un anciano. Al parecer, llevaba su familia consigo en el viaje: un sombrajo de hojas de palma en el centro del barco estaba ocupado por mujeres y niños. Al contrario de los indios desnudos del Caribe, con los que Colón se había familiarizado, las mujeres llevaban vestidos de algodón y ocultaron sus caras de las miradas españolas con chales de brillante colorido. La mercancía de la canoa comprendía muchos chales y otras prendas de algodón, como camisas sin mangas y teñidas. También había largas espadas de madera con filos de obsidiana o de sílex, campanas de cobre fundido, hachas de cobre y moldes para fundirlo. Lo que los viajeros mayas valoraban más, a juzgar por el modo como registraron el fondo de la canoa por si se habían derramado algunos, eran los granos de cacao. Los españoles tomaron muestras de todas las mercancías mayas y les dieron a cambio algunas baratijas. Los dos navíos se separaron, no sin antes que Colón se apoderara del anciano capitán para que le sirviera de intérprete.

La expansión putún había empezado a finales del siglo VIII. No fue exclusivamente económica: los mercaderes también hicieron sentir su poder militar. Sus tempranos contactos con los altos de México debieron proporcionarles una ventaja estratégica inicial sobre sus vecinos. Entre sus armas había algunos proyectiles mexicanos desconocidos para los otros mayas de su tiempo, dardos y, sobre todo, propulsores. Parece cierto que la expansión putún a lo largo de las fronteras de los llanos meridionales fue un factor de la caída de la civilización maya clásica en aquella zona. Al mismo tiempo, la expansión en la otra dirección, hacia el interior y a través de los llanos septentrionales, contribuyó a un florecimiento posterior en aquella zona maya.

En las postrimerías del siglo XII, como ha demostrado Proskouriakoff, muchos aspectos artísticos de la civilización clásica maya manifestaban un estado de decadencia que se arrastraba desde hacía mucho tiempo. En la interpretación tradicional, esta decadencia social refleja una decadencia social generalizada. Por contra, la realidad es que la vida económica de la civilización maya durante estos siglos fue ganando en complejidad y rigor.

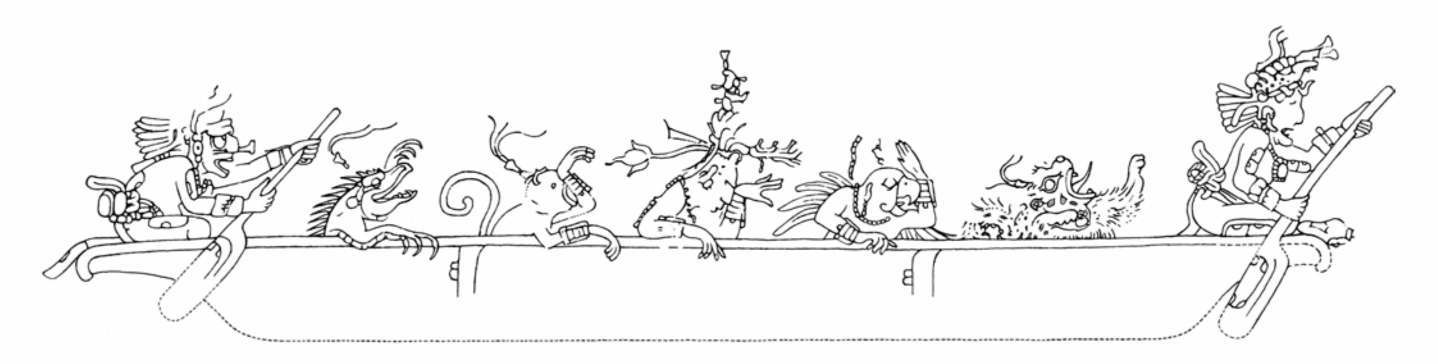
La acumulación y distribución de materias primas y objetos manufacturados se estaba expandiendo rápidamente. Por ejemplo, en la fabricación de cerámica, los métodos de producción en serie habían sustituido a la construcción singular pieza a pieza, de antaño. Al mismo tiempo, el nivel de vida de las clases populares se iba elevando; el intrincado sistema del comercio de larga irradiación dio acceso a la población rural a la gran variedad de nuevas materias que los comerciantes traían de otras áreas, cuando volvían a Yucatán. Y quizás lo más importante de todo: al descentralizarse la autoridad religiosa, la economía de mercado en pleno auge aportó otra para sustituirla, la autoridad mercantil.

En el transcurso de la expansión política putún, mercantilmente orientada hacia los llanos septentrionales de la península de Yucatán, parece que este pueblo se rigió por una ética cultural que era nueva entre los mayas: la que pudiera llamarse del pragmatismo mercantil, que reflejaba la ascendencia de una clase mercantil a expensas de la antigua teocracia.

A la hora de caracterizar el llamado período de decadencia de la civilización maya, muchos arqueólogos no han prestado atención al crecimiento de la nueva clase y al significado de su ética cultural. Los hallazgos arqueológicos de la isla de Cozumel nos proporcionan los datos para corregir este olvido.

Cozumel tiene una extensión de más de 450 km². Durante años, el Instituto Nacional Mexicano de Antropología e Historia ha ido registrando gran cantidad de yacimientos arqueológicos en Cozumel, muchos de ellos a lo largo de la costa *(véase la ilustración de la pág. 68).* Las investigaciones de A. Escalona Ramos y William T. Sanders indican que los primeros habitantes mayas llegaron a la isla a comienzos de nuestra Era. Cozumel permaneció en un tranquilo segundo orden hasta el inicio del período clásico final, alrededor del año 800 d. C. En adelante, la importancia de la isla fue en aumento; entre otros motivos, porque en ella se asentaba un santuario visitado por peregrinos de todos los rincones de Yucatán. Cozumel alcanzó su cenit de 1300 a 1500 d. C., período de la supuesta decadencia maya. Este dato de crecimiento de la isla y su preeminencia coincide de tal modo con la preeminencia de los putún que cuando, a principios de 1970, tuvimos la oportunidad de recorrer Cozumel, nos propusimos sacarle partido. Nuestras instituciones respectivas (Peabody Museum of Archaeology and Ethnology de la Universidad de Harvard y el Department of Anthropology de la Universidad de Arizona) acordaron organizar una expedición, financiada por la National Geographic Society. Trabajamos en la isla, con la autorización y cooperación del Instituto Nacional Mexicano, durante el bienio 1972-1973, como directores de un proyecto conjunto. Nuestro objetivo era ver de qué modo la isla había servido a los intereses de los comerciantes putún, y si los hallazgos arqueológicos de Cozumel reflejaban o no la cultura mercantil de los comerciantes.

En Cozumel, las estructuras arquitectónicas precolombinas de mayor volumen son de escasa categoría, y despertarían poco interés a los historiadores del arte. Sin embargo, representan una



CANOA DE SIETE PASAJEROS, que es la mayor de las representadas en el arte maya. Se encontró entre varios objetos, delicadamente labrada en hueso animal y enterrada con otros elementos funerarios en una tumba localizada debajo del templo I, en Tikal. Los remeros y sus grotescos pasajeros pertenecen todos al panteón maya; el fino dibujo, borrado en algunas de sus partes, es un ejemplo del virtuosismo del clásico temprano. La descripción de embarcaciones superiores a ésta se debe a observadores europeos, quienes afirman que en las mayores cabían cincuenta hombres.

gran inversión de trabajo y materiales por parte de los putún. La importancia principal de Cozumel, desde el punto de vista mercantil, radicaba en su situación estratégica de cabeza de puente y de almacén de materiales. Pero la isla tiene poca altitud y está sujeta a inundaciones durante la época de lluvias o de huracanes. Por ello, los putún construyeron grandes plataformas de piedra, cuya altura suficiente les permitiera hallarse seguros, lejos de la amenaza de inundaciones, en más de media docena de lugares. Las plataformas fueron construidas en mampostería irregular de piedra calcárea. La mayor está en Buena Vista; consta de un grupo de plataformas conexas, cuya altura es de cinco metros, que cubren una superficie de más de siete hectáreas.

Todas las plataformas existentes en Cozumel parecen haberse construido en el transcurso de los tres siglos del "decadente" postclásico tardío. Aunque el principal motivo de su erección debió ser proporcionar un almacén a salvo de las inundaciones, no fue esa la única razón. En efecto, ciertas mercancías putún, como la sal, llegaban al mercado estacionalmente. Si tales mercancías son almacenadas por el consumidor final, puede haber una saturación estacional, y, por tanto, se desvalorizan. Si los almacena el comerciante, usando amplios espacios como eran las plataformas de Cozumel, el precio permanecerá estable. La inversión de tiempo y esfuerzo en la construcción de las plataformas pudo resultar positiva desde más de un punto de vista mercantil.

Otro programa importante de infraestructura fue la erección de viaductos elevados; al parecer, permitían la comunicación, con independencia del clima, entre los puntos costeros para el desembarco de la mercancía y los almacenes del interior. Igual que las plataformas, los viaductos se construyeron con mampostería irregular de piedra calcárea. Hoy día no es fácil conocer su trazado, excepto en lugares donde siguen un paso seguro y seco a través de zonas permanentemente pantanosas.

Un amplio viaducto pasa a través de un arco exento, actualmente en ruinas. Ambos, viaducto y arco, parecen haber servido a otros fines distintos de los mercantiles. El arco enmarcaba, evidentemente, la entrada principal a un grupo de ruinas hoy conocidas como zona de San Gervasio. En este lugar pudo residir el famoso santuario, centro de peregrinación, del templo consagrado a la diosa Ix Chel. El viaducto pudo haber sido el camino que conducía a los peregrinos al santuario de Ix Chel, o a otro importante recinto sagrado. En el panteón maya, Ix Chel era la consorte de Itzamna; desempeñaba un papel superior al de Juno en relación con Júpiter, puesto que no sólo era diosa de la luna sino también protectora del tejido, de la



MUROS DE DELIMITACION, que dan fe del régimen parcelario vigente en Cozumel. Se les distingue como líneas blancas muy finas, en esta fotografía aérea, tomada de un lugar recién descubierto en la zona de San Gervasio, al norte de la isla. El sistema de campos amurallados sugiere la posibilidad de que los putún de Cozumel hubieran abandonado el sistema tradicional de cultivo comunal de la tierra por el de la propiedad privada.

medicina y de los partos. Las visitas a su santuario eran frecuentes, quizás porque el único oráculo de los mayas estaba localizado allí. Aunque el viaducto de San Gervasio no se destinara directamente al comercio, los putún pudieron considerarlo como una buena inversión. Quienes peregrinaban a Cozumel podían quedarse muy bien a comerciar. Por esta razón, el movimiento de peregrinos ya constituyó en sí mismo una fuente importante de ingresos para los isleños.

Un descubrimiento inesperado en la campaña de Cozumel fue el de un sistema de campos amurallados que alcanzaba a toda la isla *(véase la ilustración de la pág. 70).* Los muros se construyeron evidentemente durante los mismos tres siglos del así llamado período decadente. Como las plataformas y los viaductos, son de mampostería de piedra calcárea. Su altura varía de medio metro a metro y medio. Los muros marcaban seguramente los límites de los campos que los campesinos isleños fertilizaban quemando la broza y plantando a continuación maíz, cereales, calabazas y otras cosechas. Aunque los campos tienen un perímetro irregular, la superficie media suele ser de 4000 m².

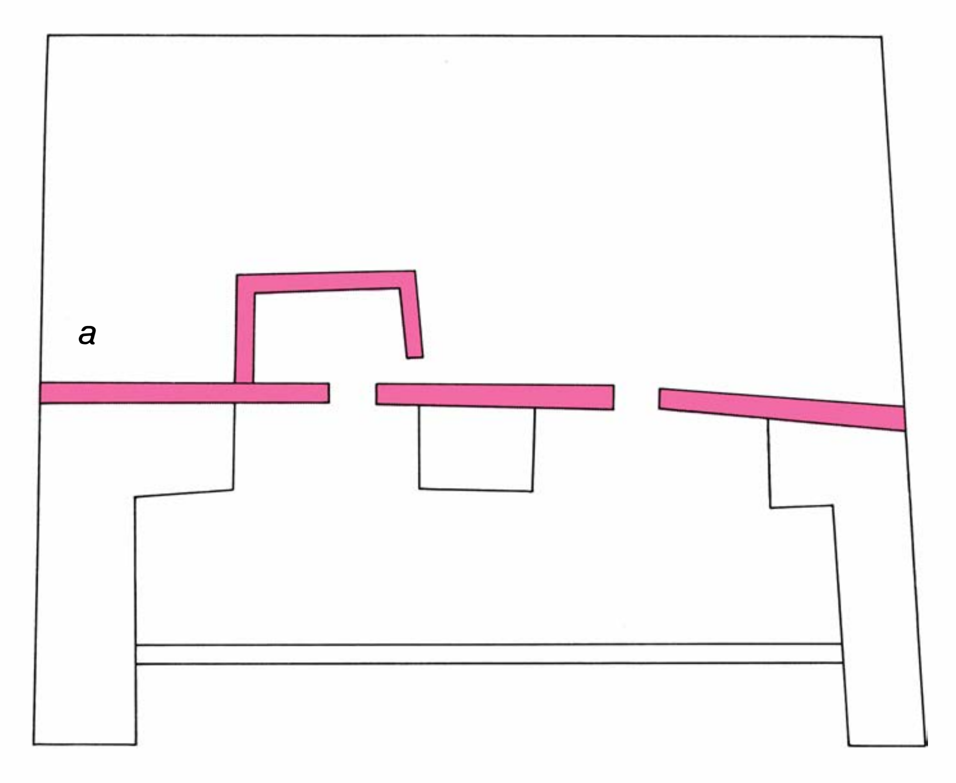
Posiblemente, la prosperidad de Cozumel, en la cima del poder de los putún, estuviera acompañada de un aumento de población suficiente, que hiciera, muy difícil la posesión de tierra en la isla. Si éste fue el caso, los muros pudieron marcar los límites de propiedad y las demarcaciones de los campos. El esfuerzo invertido en tal sistema de vallas induce a suponer que los muros representan el salto dado por los putún desde el patrón tradicional de los campos de cultivo comunales de los mayas hacia otro en que los campos eran propiedad privada.

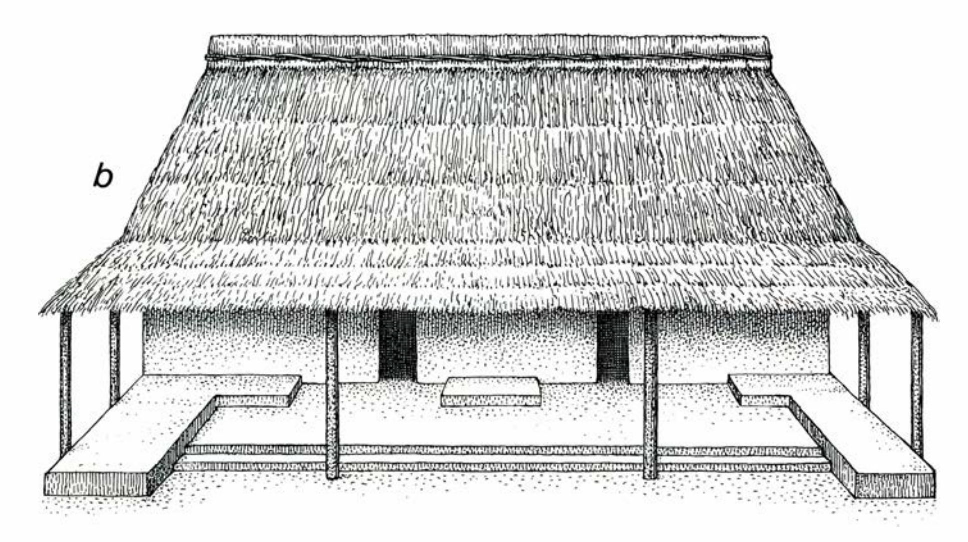
La construcción de un sistema de núcleos de defensa costera sería otra actividad putún en Cozumel, que requirió asimismo una considerable inversión en tiempo y trabajo. Todavía se mantienen en pie catorce de estas estructuras a lo largo de las costas de la isla, tanto en las de sotavento como en las de barlovento. Los relatos españoles sobre los primeros contactos con los mayas cuentan cómo hacían hogueras encima de tales estructuras, en señal de alerta a la población interior ante la proximidad de enemigos.

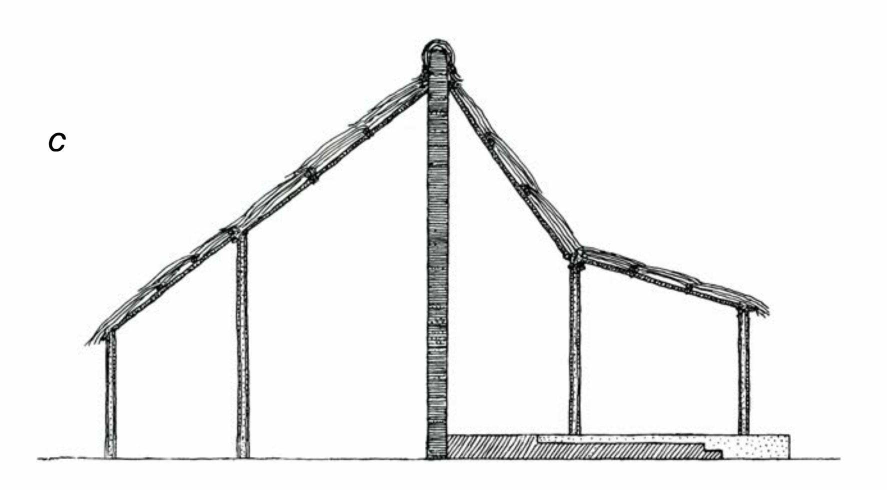
Las estructuras de esta red de alertas varían en tamaño y en estilo arquitectónico. Además de su función defensiva, algunas pudieron servir de santuarios o para otros fines. Todas, sin embargo, parecen haber sido construidas durante

INTERVALO DE TIEMPO	MEXICO	AREA MAYA	PERIODO DE COZUMEL
900 a. C. a 300 a. C.	(1250 a. C.) OLMECA (400 a. C.) / (600 a. C.)	PRECLASICO MEDIO (PRIMERA OCUPACION DE LOS LLANOS MAYAS)	(ISLA SIN OCUPAR)
300 a. C. a 300 d. C.	(50 a. C.)	PRECLASICO TARDIO (AUGE DE LOS CENTROS CEREMONIALES)	FORMATIVO TARDIO (SE OCUPA LA ISLA POR PRIMERA VEZ)
300 d. C. a 600	TEOTIHUACAN / MONTE ALBAN	CLASICO TEMPRANO (SURGIMIENTO DE LA CIVILIZACION MAYA)	PERIODO TEMPRANO I (CONTINUA LA OCUPACION)
600 a 800	(700 d. C.)	CLASICO TARDIO (PREEMINENCIA DE CIUDADES COMO TIKAL)	PERIODO TEMPRANO I (CONTINUA LA OCUPACION)
800 a 1000	(900 d. C.)	CLASICO TERMINAL (CAIDA DE LOS LLANOS MERIDIONALES)	FLORECIMIENTO MODIFICADO (AUMENTA LA PREEMINENCIA)
1000 a 1250	1000 d. C.) / TOLTECA (1190 d. C.)	POSTCLASICO TEMPRANO (PREEMINENCIA DE CHICHEN ITZA)	FLORECIMIENTO PURO (COMIENZA LA PREEMINENCIA)
1250 a 1520	(1345 d. C.) AZTECA (1521 d. C.)	POSTCLASICO TARDIO (PREEMINENCIA DE MAYAPAN HASTA 1450 APROX.)	DECADENTE (APOGEO DE COZUMEL 1400 APROX.)
1520 a 1600	CONQUISTA ESPAÑOLA Y PERIODO COLONIAL		(DISMINUCION DRASTICA DE POBLACION)

CRONOLOGIA MAYA, desde la primera ocupación de los llanos hasta la llegada de los españoles, comparada con la cronología correlativa de Cozumel *(derecha)* **y del propio México** *(izquierda).*







CASA CON FACHADA FALSA. La construcción en piedra está reducida a una pared *(a, en color).* **Constituye un ejemplo del pragmatismo putún. Todo lo que un visitante podía ver era el pórtico sombreado con un pajizo** *(b),* **con una pared finamente revocada y pintada con vivos colores. El resto de la casa era una amalgama de postes y paja** *(c).* **El plano de esta casa se dedujo a partir de unas excavaciones hechas en San Gervasio; Mayapán tiene estructuras similares.**

el mismo período de 300 años, y hay pruebas de que, en algún momento, la red incluía por lo menos veinte estructuras costeras.

Los asentamientos orientados hacia Cozumel en la tierra firme de Yucatán estuvieron fortificados durante este período. Tulum es un buen ejemplo: el acantilado sobre el que está situada la protege por la costa y el sistema de murallas de las incursiones desde tierra firme. Por contra, Cozumel carecía de defensas naturales. Más aún, los numerosos muros que rodeaban los campos en la isla no tenían la altura suficiente para protegerla. Parece que los putún se dedicaron preferentemente a construir tinglados, enclaves residenciales y estructuras religiosas y administrativas tierra adentro, a cierta distancia de la costa indefensa. Las zonas interiores se mantenían unidas a determinados puntos costeros mediante una red de viaductos. Cuando un enemigo se aproximaba, los vigías costeros podían alertar a la población de forma que las fuerzas del interior usaran los viaductos para concentrarse en el punto amenazado de la costa.

Dos hallazgos arqueológicos más en Cozumel nos ayudan a completar el cuadro de este pueblo mercader y pragmático. Primero, las casas construidas por los comerciantes tenían unas fachadas imponentes: bajo el pajizo que daba sombra al pórtico delantero de la casa, había una pared de piedra, finamente revocada y decorada con brillantes pinturas. La pared estaba recortada por las puertas que conducían al interior sombreado *(véase la ilustración de la izquierda).* Este muro, sin embargo, era la única construcción en piedra de la casa. Los portalones conducían directamente a un interior techado con paja, de casi la misma superficie que el pórtico cubierto, y cerrado por tres paredes hechas de postes de madera. Estas engañosas fachadas de piedra, parecidas a las construcciones de cartón-piedra de Hollywood, se hicieron no sólo en Cozumel, sino también en los otros enclaves de Yucatán dominados por los putún.

El segundo hallazgo fue el descubrimiento de varios depósitos de hojas de hacha, talladas en jadeíta verde, que los mayas preferían para sus objetos ceremoniales. Cuando se descubre un depósito de éstos en un yacimiento del período clásico, la talla del jade suele ser algo admirable. Sin embargo, en depósitos de Cozumel, pertenecientes al clásico tardío, la calidad es mínima. Casi todas las hojas de hacha que se han en-

contrado están rotas o con muescas. Se saca la impresión de que se trata de mercancías rebajadas, de segunda clase e irregulares.

¿Qué conclusiones pueden deducirse de esta descripción de una clase mercantil y pragmática? Hemos construido un modelo basado en la interpretación de los resultados obtenidos en nuestra campaña en Cozumel y en una reinterpretación de los hallazgos conseguidos en distintos lugares de Yucatán por otros investigadores. A grandes rasgos, el modelo podría ser cierto para una clase mercantil pujante en otras muchas sociedades distintas, pero en sus detalles es peculiarmente maya. Su corolario principal es que las evoluciones de los tres siglos así llamados decadentes, tomados tradicionalmente como signos de decadencia social, aparecen lógicos y progresivos cuando se contemplan desde un punto de vista distinto del de la élite. En otras palabras, se tiene un cuadro diferente de la civilización maya cuando se la considera desde la base que cuando se la ve desde arriba.

El dominio putún de Yucatán es una prueba de que la élite teocrática anterior había cedido su puesto a una clase mercantil. Lo mismo que otros comerciantes a lo largo de la historia, los putún querían conservar su capital líquido. Gastarse los beneficios del comercio en lujosos refinamientos para las tumbas no tenían utilidad para ellos. El jade imperfecto sirve a los muertos tan bien como el otro. Siguiendo el mismo razonamiento, la erección de grandes monumentos daba pocos beneficios al comercio; un almacén seco es más útil que un templo adornado. Sin embargo, las apariencias mantenían su lugar destacado. Una fachada barata, revocada con una capa gruesa que ocultara las piedras desiguales, era una inversión lógica. Tal muro no tenía por qué durar perennemente, pero los putún no construían para la posteridad.

Unos cuantos ejemplos más destacan el pragmatismo de los comerciantes de Cozumel, hombres cuya supervivencia dependía de su pronta reacción ante las circunstancias políticas y económicas cambiantes. Un cambio de estas características ocurrió con la caída del poderío tolteca en el siglo XIII. Cuando la principal ciudad tolteca en Yucatán, Chichén Itzá, fue repentinamente abandonada, toda la red de comercio en dicha península perdió su plaza central. Para restablecer los enlaces perdidos se fundó en pocos años un nuevo centro, Mayapán. Un hecho que apoya la tesis de que la nueva ciudad se construyó para tomar el relevo de la antigua es que los primeros edificios de Mayapán reproducen casi exactamente el centro de Chichén Itzá, a escala menor. Hay indicios en la literatura arqueológica de que los fundadores de Mayapán llegaron de Cozumel con pleno apoyo de los gobernantes de la isla. Tal acción hubiera sido lógica; a los putún de Cozumel les interesaba estabilizar la red comercial desbaratada en Yucatán, tan pronto como fuera posible.

Un cambio todavía más drástico acompañó a los primeros viajes de incursión en Yucatán desde Cuba por parte de los españoles. Los tres barcos de Córdoba en 1517, los cuatro de Grijalva en 1518 y la flota de Cortés en 1519. Los mayas de tierra firme repelieron a los hombres de Córdoba cuando éstos tomaron tierra en la costa septentrional de Yucatán y mataron a la mayoría de ellos cuando atracaron en la costa occidental. Grijalva tomó tierra por dos veces en Cozumel sin ningún incidente: los isleños se retiraron al interior al llegar los españoles. Cuando Grijalva mandó a un jamaicano bilingüe, que vivía en Cozumel, para citar a los jefes mayas a una reunión, éstos se negaron a contestar. Este comportamiento pasivo contrasta rotundamente con la reacción hostil de los mayas del continente a las incursiones de Grijalva.

Cortés y su flota se encontraron en Cozumel con una respuesta completamente distinta, si damos crédito al relato de un testigo ocular como Bernal Díaz. Por casualidad un barco (en el que Díaz navegaba) llegó a la isla adelantándose al resto de la flotilla. Como cuando Grijalva llegó a Cozumel, los habitantes huyeron al interior. En un asentamiento vacío, los españoles que habían tomado tierra se apropiaron de cuarenta volátiles (seguramente pavos) y del tesoro de un santuario, que guardaba, entre otros objetos, pendientes de oro con una alta aleación de cobre. También aprehendieron a tres putún que se habían rezagado. Cortés llegó poco después, y ordenó que soltaran a los prisioneros. Les devolvió todo lo que se había robado del santuario (los volátiles se los habían comido ya), les agregó incluso algunos productos comerciales españoles como regalo y les rogó que buscaran a sus autoridades y prepararan un encuentro. Al día siguiente, el oficial en jefe del asentamiento se presentó, seguido de los hombres, mujeres y niños que habían huido. Estuvieron entre nosotros, cuenta Díaz, como si hubieran sido amigos nuestros de toda la vida.

Cortés sabía que cuando Córdoba tomó tierra en la costa occidental de Yucatán en 1517, los mayas, al ver a los españoles, les llamaron *"castilan"*. Para un buen vasallo de Castilla y León, se trataba de demasiada coincidencia para no tenerla en cuenta, y Cortés preguntó a las autoridades de Cozumel si sabían de otros españoles en la zona. Las autoridades le constestaron que, efectivamente, dos españoles vivían en tierra firme y que algunos comerciantes habían hablado con ellos pocos días antes.

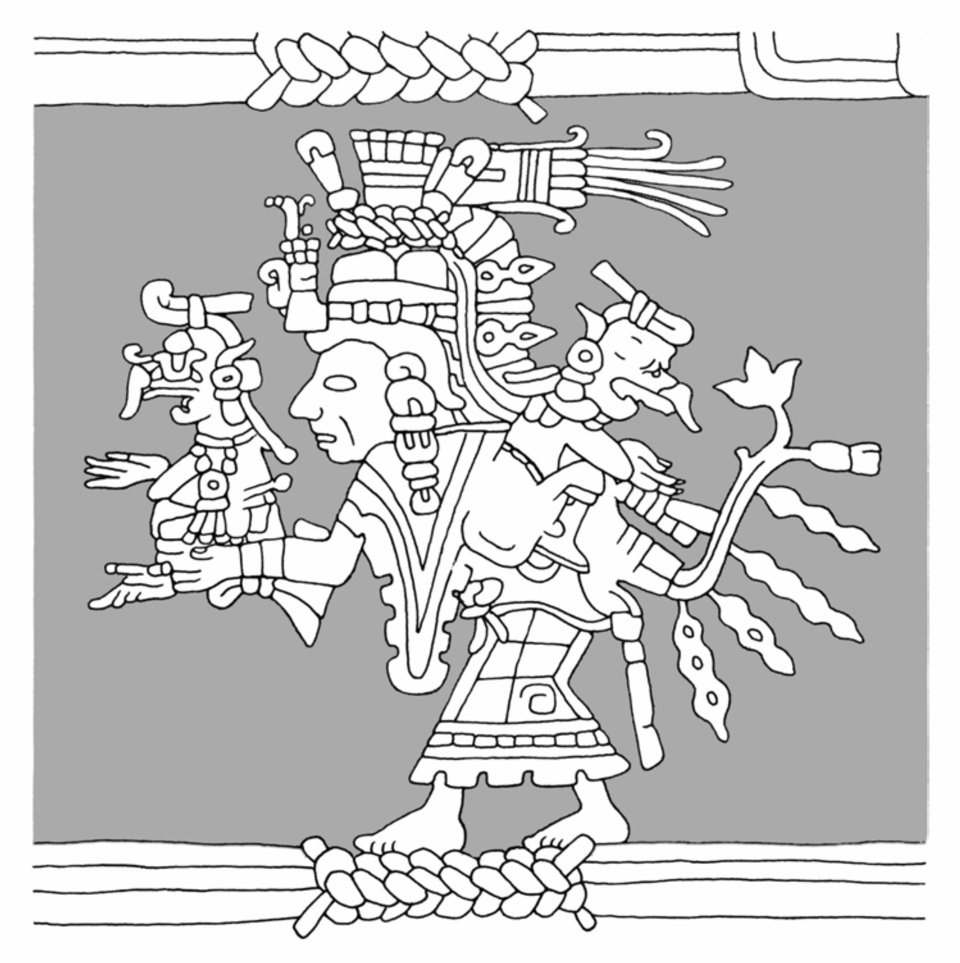
Cortés pidió a los comerciantes que les llevaran cartas a sus desconocidos paisanos, junto con algunos bienes para el caso de que tuvieran que ser comprados a los mayas que los retenían. Mandó entonces a los comerciantes a tierra firme con uno de sus navíos. El resultado fue que rescataron de la esclavitud a Gerónimo de Aguilar, que se había quedado en la costa de Yucatán en 1511 durante un viaje de Panamá a Santo Domingo. Aguilar, que hablaba ya muy bien el maya, se unió a Cortés para mediar de intérprete (el otro español, un compañero de barco de Aguilar, se quedó en Yucatán, con su mujer maya y sus niños).

Poco después, Cortés profanó y descralizó lo que parece haber sido el santuario de Ix Chel. Los españoles no sabían nada de la diosa, pero Díaz afirma que al templo acudían muchos peregrinos de tierra firme. Más aún, de hecho presenció cómo lo que pudo haber sido el oráculo parlante (un viejo indio vestido con larga túnica) profería lo que Bernal Díaz calificó de sermón negro. Cortés oyó también al oráculo y, después de manifestar ardientemente a las autoridades de la isla que no podía venir nada bueno de la idolatría, él y sus hombres destruyeron las imágenes del templo y ordenaron a los constructores mayas que erigieran un nuevo altar. Cuando estuvo terminado, Cortés puso en él una imagen de la Virgen María y una cruz de madera. Sus actos no provocaron ningún reto armado por parte de la población de Cozumel, ni que ésta huyera. No llegaron a alterarse, e incluso quemaron incienso.

¿Cómo debe interpretarse el que los mayas de tierra firme tuviesen un comportamiento hostil y neutral los mayas de Cozumel? Por una parte, como comerciantes que recorrían largas distancias, los putún estaban acostumbrados a encontrarse con extranjeros. Si creían que las intenciones de estos extraños eran hostiles, que es lo que ocurrió cuando vieron a Grijalva o a los hombres del primer barco de Cortés, hacían todo lo posible para evitar su contacto. Desde



LA FACILIDAD DEL TRANSPORTE aconsejaba la uniformidad del diámetro y la inclinación de las paredes de estos cuencos mayas de poca profundidad; así se podían embarcar apilados. Fueron hallados en Cozumel. Quizás son mercancías de los putún traídas del Yucatán septentrional.



LA DIOSA MAYA Ix Chel aparece representada sosteniendo dos imágenes del dios de la lluvia Chac, en esta pintura mural de un templo de Tulum, yacimiento del postclásico tardío en tierra firme. El santuario sito en Cozumel albergaba al único oráculo conocido en el antiguo mundo maya.

hacía tiempo, y siguiendo las rutas comerciales, debían haber llegado a sus oídos las matanzas que los españoles podían llevar a cabo con sus armas extrañas y temibles.

Quizás, también, los putún se dieron cuenta del potencial político y económico de los españoles. Está claro que prefirieron no tomar en consideración las acciones de Cortés en el templo de Ix Chel. Su información sobre los españoles que vivían en tierra firme y la entrega de los mensajes de Cortés fueron también actos de buena voluntad. Hay pruebas, en la despedida que hicieron a Cortés, de que los pragmáticos putún vieron en los españoles la configuración del futuro. Antes de que Cortés se embarcara, los putún pidieron y recibieron una "carta de recomendación" para enseñarla a otros viajeros españoles que llegaran a Cozumel y de este modo asegurarse el buen trato por parte de éstos.

Como ejemplo final del pragmatismo de los putún, consideremos el santuario de Ix Chel y su oráculo. ¿Puede ser una pura coincidencia que el único santuario con oráculo en el mundo maya estuviera en Cozumel? Los oráculos se hallan en una posición de privilegio que les permite una mayor flexibilidad ante los cambios sociales y culturales respecto a los sacerdotes conservadores que siguen un ritual en otra clase de santuarios. Hay incluso una muestra de este tipo de flexibilidad en la historia moderna de los mayas.

Durante la larga y sangrienta "guerra de las castas", que destrozó periódicamente Yucatán desde 1847 a 1901, los rebeldes mayas acudían regularmente a pedir consejo a Chan Santa Cruz, un santuario existente en el fuerte de Quintana Roo. Allí, los dioses, a través de los sacerdotes, hablaban por medio de las tres "cruces parlantes", que dictaban la política bélica a los rebeldes.

Y ahora ya podemos formularnos una última pregunta. ¿Puede considerarse, en conciencia, la última fase decadente de una civilización moribunda un período de la historia que contempla el surgimiento de una clase mercantil, el desarrollo de una nueva ética y un aumento sustancial en el entramado económico (que comprende acontecimientos tales como la introducción de la fabricación en serie y la mejora del nivel de vida)? Viendo el proceso desde arriba, desde una perspectiva elitista, la respuesta sería tal vez afirmativa. Viendo el proceso desde la base, creemos que la respuesta será claramente negativa. A la hora de examinar el pasado no deberíamos limitarnos a un solo punto de vista social.

así ven INVESTIGACION Y CIENCIA

«Hace cierto tiempo, pensando en la necesidad de publicar una buena revista de divulgación científica en España, consulté con mi amigo y colega, el bioquímico Alberto Sols, acerca del mejor enfoque a tomar en tal proyecto. Alberto me contestó que lo más práctico sería traducir al español la mejor revista que existe en este campo, es decir, SCIENTIFIC AMERICAN. Afortunadamente, gracias a la Editorial Labor, esto ya es una realidad. Y también un gran acierto, porque la supervivencia del hombre en la tierra en los próximos siglos dependerá en gran parte de que éste reciba una buena formación científica y humanística.»



JUAN ORO

Profesor de la Universidad de Houston, director científico del Proyecto Vikingo de la NASA.



«Me complace enormemente la noticia de la publicación de INVESTIGACION Y CIENCIA, edición en español de SCIENTIFIC AMERICAN, la revista de mayor prestigio en el mundo de la ciencia, sin otro norte que el avance de ésta y su incidencia en la vida del hombre. Qué duda cabe que ha de ser un vigoroso instrumento de progreso para los países de habla española.»

JOSEP TRUETA

Profesor Emérito de Cirugía de la Universidad de Oxford.



«Pienso que acaso la más radical de las deficiencias históricas del alma española sea la escasez de su formación científica, la carencia o la debilidad de los hábitos mentales en cuya virtud es posible hacer ciencia propia y estimar adecuadamente la que los demás han hecho. Mídase, pues, según esto, la importancia que atribuyo a la publicación de la revista INVESTIGACION Y CIENCIA. En ella veo un eficaz recurso para acercarnos al siglo XXI a la altura que éste va exigir a todos los hombres.»

PEDRO LAIN ENTRALGO

Catedrático de la Universidad de Madrid. Director del Instituto Arnau de Vilanova del Consejo Superior de Investigaciones Científicas.





«Soy lector, desde hace muchos años, de SCIENTIFIC AMERICAN, que es modelo en el difícil arte de la divulgación científica de altura. Por ello, considero que INVESTIGACION Y CIENCIA contribuirá a crear el clima social que necesita el desarrollo de la investigación entre nosotros, dando a conocer además la gran tarea ya realizada por los científicos españoles e hispanoamericanos.»

EDUARDO PRIMO YUFERA

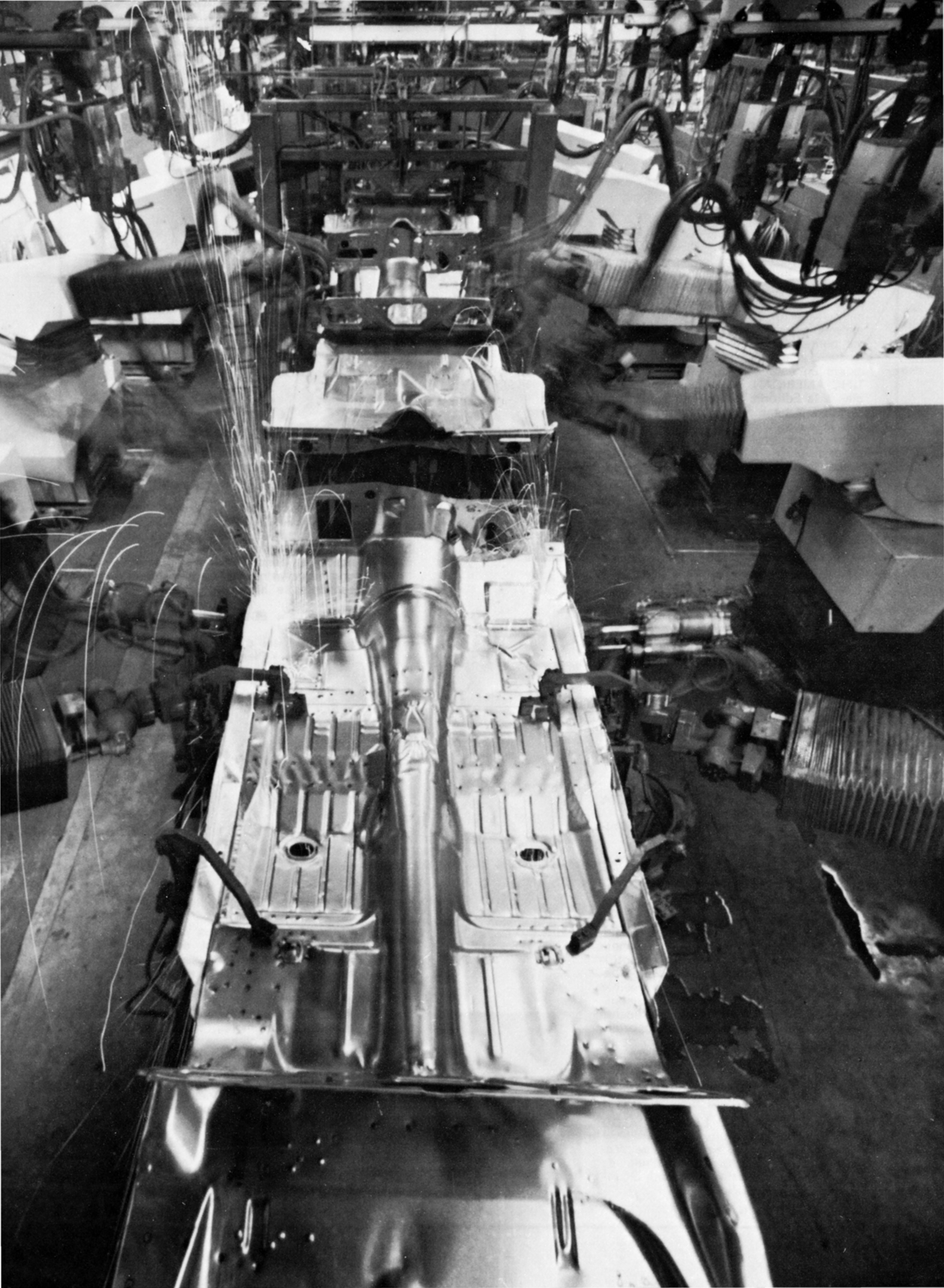
Presidente del Consejo Superior de Investigaciones Científicas.



«Tengo muchas razones para creer que la edición en español de SCIENTIFIC AMERICAN va a representar un acontecimiento de real importancia en la vida intelectual española. Entre otras, el uso continuo que de SCIENTIFIC AMERICAN hacemos la mayoría de los que, sabiendo inglés, trabajamos en el campo de las ciencias humanas. La calidad y claridad de los trabajos, así como su amplia y actual temática, hacen de esa publicación una herramienta de trabajo en verdad indispensable, que ahora estará al alcance de todo el mundo.»

JOSE LUIS PINILLOS

Catedrático de Psicología de la Universidad Complutense, Madrid.



Sistemas robot

Muchas fábricas empiezan a utilizar máquinas capaces de recibir instrucciones y de realizar tareas humanas sencillas. Se está preparando a algunas para que puedan afrontar indeterminaciones de su entorno

James S. Albus y John M. Evans, Jr.

La fascinación del hombre por las máquinas que se mueven por sí mismas es tan antigua como la historia humana. Se dice que, en la Grecia de Aristóteles, muchas escenas de ciertas representaciones corrían a cargo, exclusivamente, de marionetas automáticas que se accionaban mediante unos pesos que colgaban de cuerdas arrolladas. Muchos siglos más tarde, las familias reales europeas se aficionaron a unos autómatas de forma humana que podían escribir, dibujar y tocar instrumentos musicales. En los últimos años se ha desvanecido gran parte del aura mágica que rodeaba a los autómatas mecánicos. Hoy en día, y por todas partes, se están utilizando máquinas automáticas y robots industriales para efectuar trabajos que comportan un riesgo grave, son pesados o, sencillamente, tienen un rendimiento económico negativo cuando los realiza directamente el hombre.

El uso, en la industria, de los mecanismos de control automático no es nuevo, por supuesto. Hace aproximadamente doscientos años que James Watt inventó el regulador centrífugo que controla la velocidad de las máquinas de vapor. Fue la primera vez que se aplicó el principio de realimentación para detectar y corregir las desviaciones producidas respecto a los valores deseados de una señal de control. Durante los años treinta y cuarenta de nuestro siglo, las refinerías de petróleo y las plantas petroquímicas se "robotizaron" ampliamente con la introducción de instrumentos analógicos de control, sencillos, en los bucles de realimentación, éstos regulaban la presión, la temperatura y la velocidad del flujo en las columnas de destilación, en los *crackers* catalíticos y en otros equipos destinados al tratamiento de los materiales de flujo continuo. Durante la Segunda Guerra Mundial el Laboratorio de Servomecanismos del Instituto de Tecnología de Massachusetts desarrolló las bases matemáticas del control por realimentación, necesario en piezas de artillería que usaban el radar para hacer puntería sobre blancos de gran velocidad. Nacidas en las postrimerías de los años cuarenta, las técnicas de fabricación automática se extendieron a las industrias del tratamiento y transformación del hierro, gracias a la tecnología electrónica, que había cobrado auge durante la Segunda Guerra Mundial. El desarrollo de los ordenadores digitales aceleró dicho proceso de expansión. En los años cincuenta entró en funcionamiento maquinaria auxiliar controlada por números perforados en cintas de papel; en la década siguiente se comercializaron los primeros robots industriales, controlados ya por ordenadores electrónicos que se diseñaban específicamente a tal fin.

Con la aplicación de los ordenadores electrónicos a las máquinas industriales, los robots se han convertido en un tema del mayor interés científico y técnico. La inserción de un ordenador en el bucle de realimentación permite un repertorio de estrategias de control mucho más amplio que el que podía conseguirse con los sencillos bucles de realimentación utilizados anteriormente. La facultad de calcular con gran rapidez y el poder de decisión del ordenador permiten cambiar los programas de control del robot cuando sea conveniente y, posiblemente, reaccionar ante un entorno aleatorio con cierta independencia de juicio.

Hoy, la tecnología de los robots tiene dos vertientes principales, una tecnológica y la otra científica. En el desarrollo de los robots industriales prácticos, los criterios que priman son la fiabilidad y la rentabilidad. Por su parte, el estudio científico de las posibilidades del robot (encabezado a menudo por el rótulo de "inteligencia artificial") se centra, principalmente, en los problemas fundamentales de la percepción sensorial, el control de la fuerza motriz y el comportamiento inteligente.

En las fábricas de todo el mundo hay varios miles de artificios impropiamente llamados robots. La mayoría de estos mecanismos presentan pocas de las características que cualquier individuo asociaría al término robot; suele tratarse de máquinas de "manipulación", capaces de realizar únicamente los movimientos más simples. Tienen poco o ningún poder de detectar las condiciones de su entorno; su funcionamiento se reduce a ejecutar una serie de operaciones programada de antemano. Los límites de la movilidad de cada articulación de la máquina vienen fijados por topes mecánicos y cada fase del movimiento tiene que regirse por un impulso eléctrico o neumático, originado en un tablero de man-

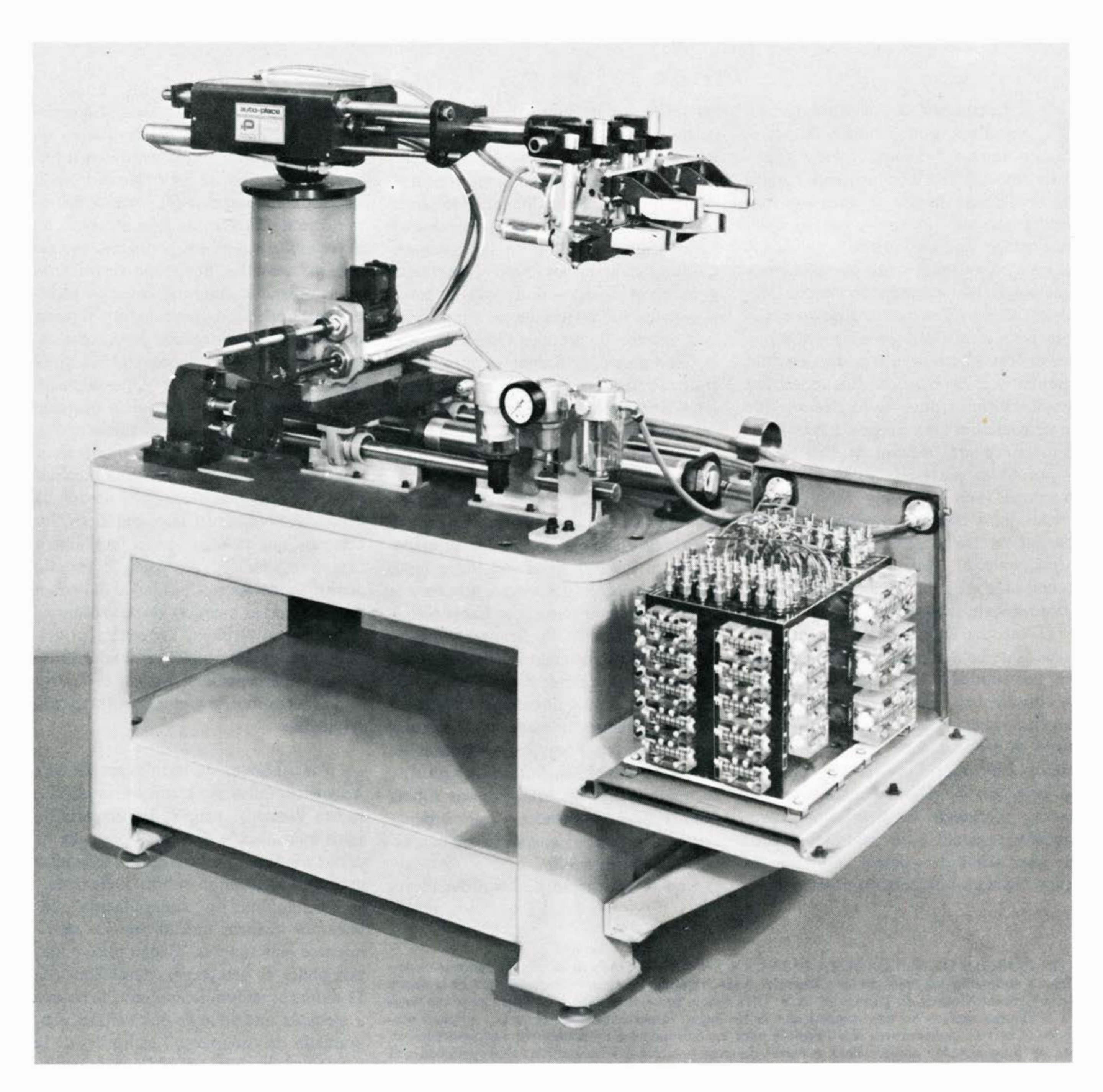
LOS ROBOTS DE SOLDADURA ELECTRICA POR PUNTOS de la página anterior se destinan a ensamblar los bajos de los Chevrolet Nova, Pontiac Ventura y Buick Skylark en la planta de la General Motors, de Tarrytown, N.Y. Diez robots denominados Unimate (que no se ven todos en la figura) sueldan los tres componentes de los bajos: la parte anterior, el suelo y la parte posterior. Los tres componentes son idénticos para los tres modelos de automóvil; van montados sobre un transportador que los lleva a través de cinco estaciones de soldadura. Cada estación está servida por dos robots Unimate. El transportador se detiene algo menos de un minuto en cada estación, tiempo en el que cada robot efectúa de 17 a 30 soldaduras. Las soldaduras tienen un diámetro de unos 13 milímetros y están separadas unas de otras por unos 5 centímetros. Los robots reemplazan las prensas de soldadura, grandes máquinas provistas de múltiples pistolas de soldar que tenían que sufrir una modificación radical cada vez que cambiaba el diseño de las piezas a soldar. Ahora, si cambia el diseño, basta con "adiestrar" de nuevo a los robots *(véase pág. 79)*.

dos formado por un conjunto de clavijas.

En un caso típico, cada fila del tablero representa un grado de libertad de una articulación específica del robot y cada columna representa una fase del programa. Las filas y las columnas se conectan convenientemente para determinar, en cada fase, qué articulación debe excitarse y en qué dirección. Cuando hay que cambiar de programa, se colocan los topes en nuevas posiciones y se modifican las conexiones de las clavijas para obtener una nueva serie de movimientos. Se puede obtener un nivel de control más complejo con la incorporación de servomecanismos que permitan asumir cualquier valor a cada grado de libertad. La adición del servocontrol exige que sensores tales como potenciómetros, codificadores y resolutores, que miden la posición de cada artículo, proporcionen realimentación. Las posiciones medidas se comparan con las posiciones previstas, y todas las diferencias que pudieran darse se corrigen por señales que se envían a los actuadores de las articulaciones correspondientes.

Hay varias formas de programar un manipulador servocontrolado. Desde el punto de vista teórico, el más sencillo consiste en un tablero de potenciómetros semejante al tablero de las clavijas. En cada fase, las posiciones de las articulaciones vienen determinadas por los valores a que están ajustados los potenciómetros, y no por los topes mecánicos.

La capacidad de los robots movidos por servomecanismos aumenta mucho si se les añade una memoria electrónica y una circuitería digital de control. El tipo de robot servocontrolado más po-



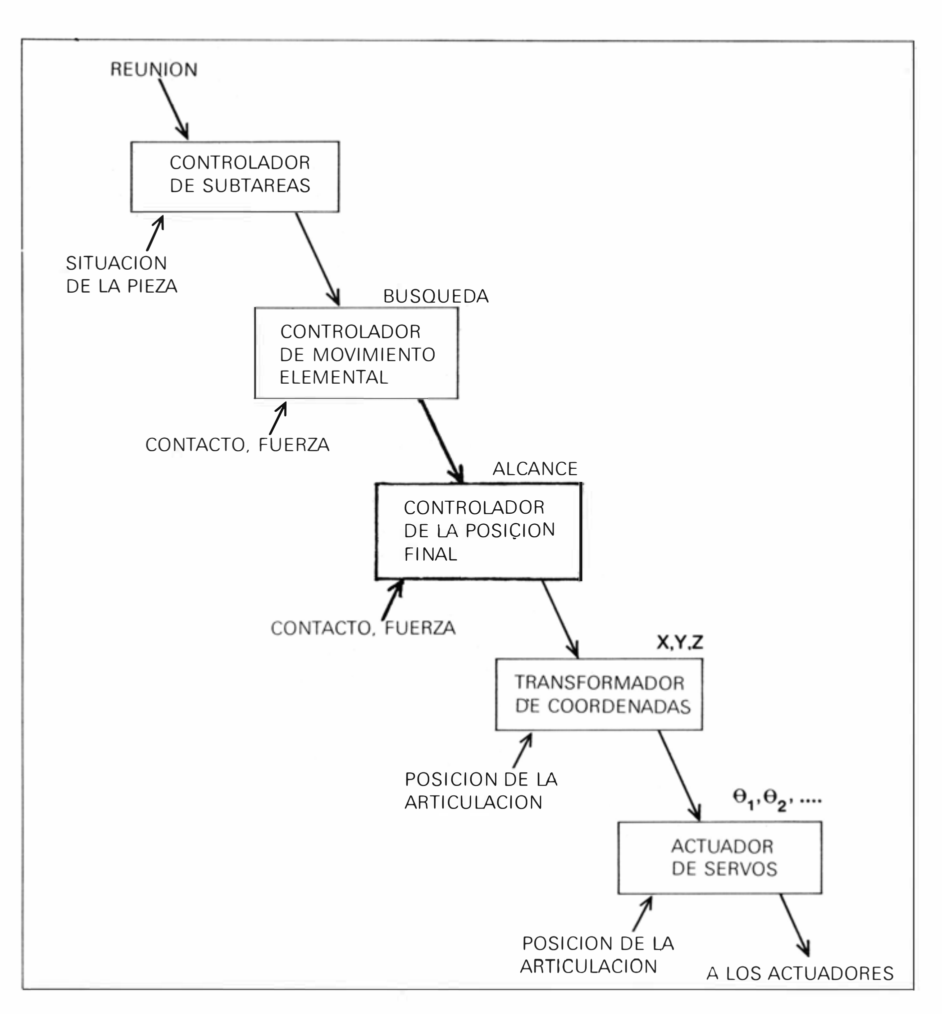
LOS MANIPULADORES que toman y colocan objetos son los robots más simples que existen en la actualidad. Este modelo, construido por la Auto-Place Incorporated, de Troy, Michigan, se acciona por seis cilindros de aire de doble acción, que permiten al robot levantar, alcanzar, girar y asir objetos, e incluso darles la vuelta. La secuencia de las operaciones, programada por un módulo de secuencia, viene activada por un módulo de potencia; ambos en la parte inferior derecha. El modelo levanta objetos de 144 kg de peso y desarrolla fuerzas de agarre de 103 kg.

pular empleado en la industria norteamericana utiliza servomecanismos digitales y tiene una memoria de placa de alambre para almacenar los programas. Para programar al robot se le hace seguir la ejecución de una determinada secuencia de posiciones. El operador utiliza una caja de control manual con mandos para controlar cada una de las articulaciones del robot. Manejando estos mandos el operador conduce al robot a la posición deseada en cada fase del programa. Apretando un botón se registra la posición de cada articulación. En el momento de reproducir el programa el sistema de control ordena, sencillamente, a cada articulación que se mueva hacia la posición registrada anteriormente para cada fase. El robot, en cuanto ha entrado en funcionamiento dentro de la cadena de producción, repite una y otra vez el programa registrado, pasando de una fase a la siguiente según un ciclo establecido, a saber: por haber quedado terminada la última fase o en respuesta a una señal de sincronismo procedente de otras máquinas.

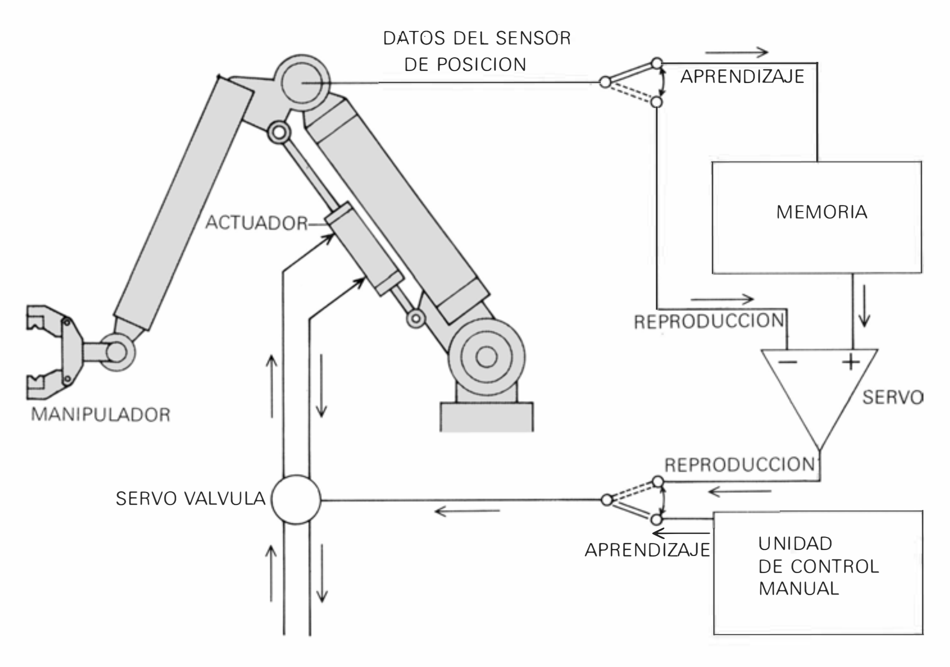
Una memoria electrónica permite al robot almacenar diversos programas y seleccionar uno u otro según sean las señales de control o las de realimentación que proceden de los sensores externos. Por ejemplo, los robots que sueldan las distintas partes de un automóvil pueden programarse para habérselas con modelos diferentes de automóviles, que vengan mezclados en la cadena de montaje.

La mayoría de los robots tienen memorias cuya capacidad de almacenaje no supera unos cuantos centenares de fases de programa. En cada fase, los movimientos correspondientes son amplios y bruscos. Si se requiere un movimiento suave y continuo, puede añadirse al sistema de control una cinta magnética capaz de almacenar muchos miles de fases muy próximas unas de otras. Cuando se quiera que la "mano" del robot se mueva suavemente en el espacio, se pondrán en juego programas de trayectoria continua como, por ejemplo, en soldaduras por arco, pintura por pulverización o recubrimiento con fibra de vidrio.

La adición de un ordenador, dotado de unidades aritméticas y lógicas, al sistema de control de un robot, permite lograr niveles de control más elevados. Con un ordenador se puede programar el robot para que mueva su mano a lo largo de trayectorias rectilíneas o de otras curvas geométricas, usando un sistema de coordenadas externo (en vez del meramente interno). Este sistema tie-



LA JERARQUIA DE LOS NIVELES DE CONTROL proporciona un método eficaz para organizar el flujo de las instrucciones y de las informaciones de realimentación, que el robot necesita para realizar un trabajo determinado. En cada nivel de la jerarquía, las órdenes del nivel inmediato superior se descomponen en órdenes más sencillas que luego se envían al nivel inmediato inferior. Informaciones procedentes de la realimentación, como son la fuerza o la posición de una articulación, pueden modificar las órdenes emitidas a ciertos niveles. En el nivel ínfimo, se envían señales de tensión de mando a los motores o a las válvulas que accionan las articulaciones del manipulador.



LA PROGRAMACION DE UN ROBOT se hace por un controlador manual, el cual guía al robot a través de una serie de operaciones. Las posiciones consecutivas de las articulaciones se almacenan en una memoria. Al pasar del modo "aprendizaje" al modo "reproducción", se repiten posiciones memorizadas. El robot está ya listo para ocupar su lugar en la cadena de producción.

ne muchas ventajas cuando al robot se le asigna la misión de colocar objetos en formaciones regulares sobre plataformas, o cuando debe tratar con objetos que se deslizan sobre la cinta transportadora de una cadena de montaje. El ordenador hace posible que la definición matemática de la trayectoria deseada, en un sistema de coordenadas del espacio de trabajo, se transforme en una señal de mando para los servomecanismos de las articulaciones de los brazos de los robots.

Daniel E. Whitney, del Charles Stark Draper Laboratory, ha desarrollado un programa de ordenador (conocido como "programa de control de posicionamiento y velocidad") para realizar esta transformación de coordenadas. Whitney utiliza un sistema de coordenadas solidario con la mano del robot. Las órdenes pueden darse en términos del movimiento de la mano, tal como "desplazarse y unidades a lo largo del eje de la mano" o "mover x unidades a la derecha". Un robot industrial desarrollado recientemente, y que tiene una facultad parecida para manejar las coordenadas del espacio de trabajo, realiza las transformaciones necesarias en tiempo real durante la ejecución del programa. El sistema de control por ordenador de este robot calcula asimismo las curvas de aceleración y deceleración que hacen máxima la velocidad a que pueden manipularse los objetos pesados.

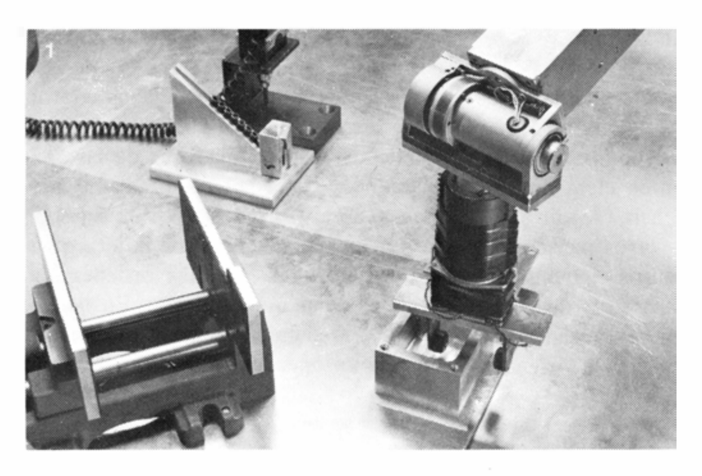
UN SENSOR DE PROXIMIDAD sujeto a las pinzas del robot permite a éste localizar y asir objetos cuya posición no se mantiene con exacta regularidad. La radiación infrarroja emitida por unos diodos luminosos, y transportada por haces de fibras ópticas, queda proyectada hacia abajo en forma de dos rayos. La radiación que refleja el objeto penetra en una serie de haces de fibras ópticas paralelas. La intensidad de la radiación reflejada actúa como una señal de realimentación que informa al robot de la distancia a que se halla la pinza del objetivo. El sistema experimental de la fotografía se halla, en el momento actual, en fase de desarrollo en el National Bureau of Standards.

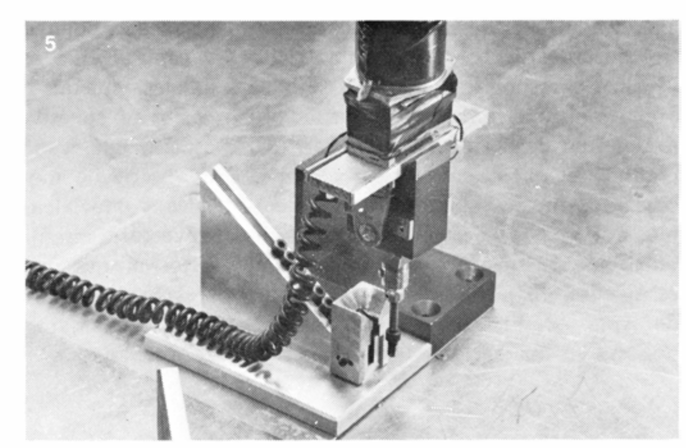
La gran mayoría de los robots industriales sólo disponen de las facultades de manipulación más elementales. La realimentación se limita a la información sobre la posición de las articulaciones, combinada con unas pocas señales de sincronismo. Por tanto, la mayoría de los robots pueden funcionar únicamente en un medio en donde los objetos que hay que manipular estén situados siempre en la posición adecuada para que el robot los pueda asir.

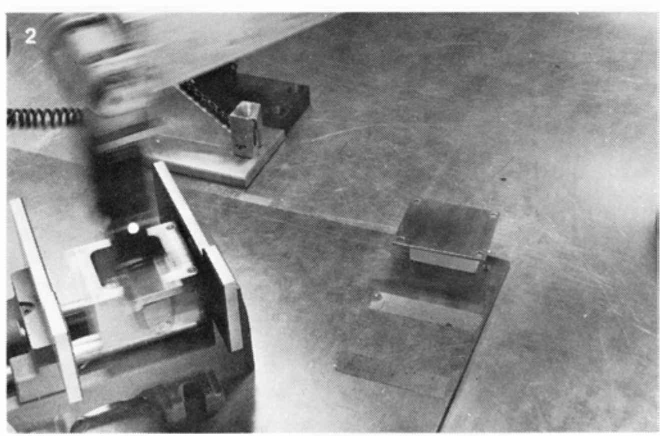
En muchas aplicaciones industriales, este nivel de ejecución basta, para llevar a cabo gran variedad de trabajos. Por ejemplo, los ingenieros de los laboratorios Kawasaki, en Japón, han demostrado que los robots pueden ensamblar piezas complicadas de motores y cajas de cambio sin más que utilizar una realimentación de posición de alta precisión, junto con pinzas bien diseñadas y fijaciones que puedan ceder ligeramente cuando se acoplan las partes. Otros experimentos han demostrado que una pequeña vibración o vaivén, juntamente con ciertos machos y ranuras diseñados convenientemente, pueden superar ligeros desajustes y evitar que se produzca un atasco cuando se ensamblan piezas de pequeña tolerancia.

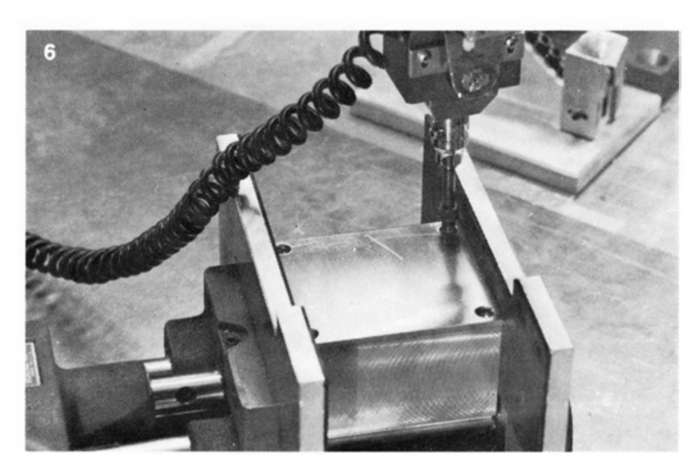
En otros muchos casos, se ha visto que resulta extremadamente caro hacer que el lugar de actuación del robot se ajuste a las condiciones imprescindibles para su correcto funcionamiento. Así se ha escrito, por ejemplo, que el equipo necesario para mantener los automóviles en la posición exacta que exige el robot encargado de la soldadura cuesta unas diez Veces más que los tales robots. Por cuya razón, algunos laboratorios estatales, universitarios e industriales están estudiando nuevos métodos de realimentación que sean capaces de compensar el hecho de que no se conozca con exactitud la posición de las piezas de trabajo.

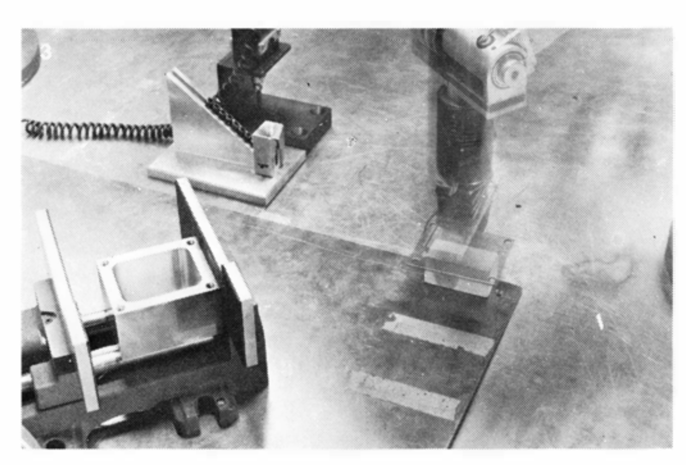
A fin de obtener un marco unitario para tratar la tecnología de los robots, el problema del control puede considerarse como una jerarquía de subproble-

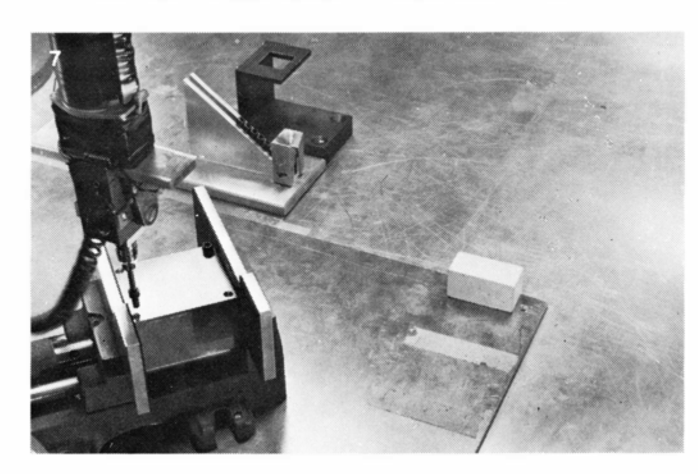


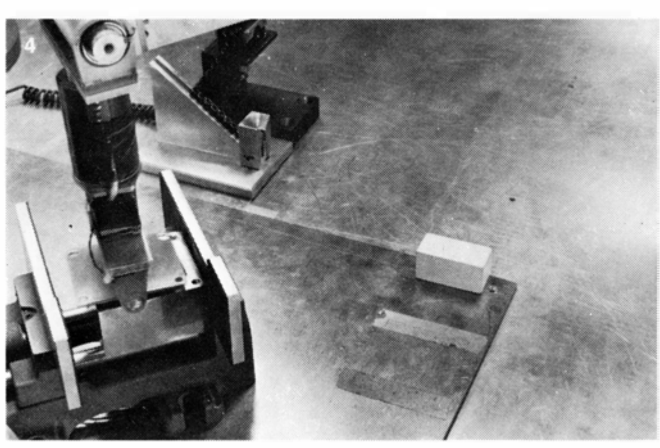


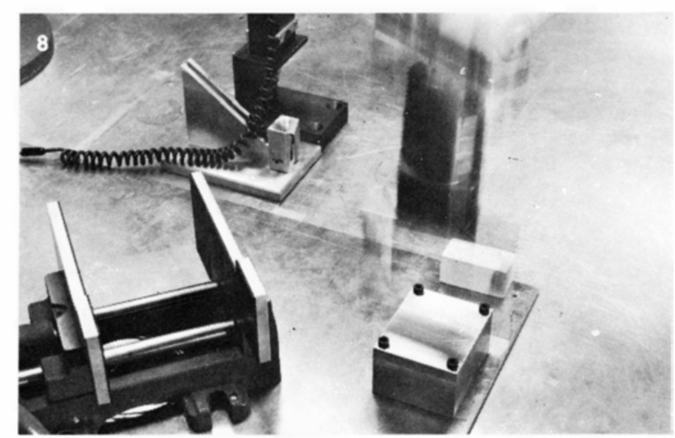












OPERACIONES DE ENSAMBLAJE POR MEDIO DE ROBOTS, ensayadas en el laboratorio de inteligencia artificial de la Universidad de Stanford. En la demostración que aparece en estas ocho fotografías, el robot ubica en primer lugar una caja de acero (1), la coloca luego sobre un torno (2). El robot regresa en busca de una cubierta (3) y la pone encima de la caja (4). El torno se cierra. El robot toma ahora una llave de contacto especial y saca un perno de la cajita (5). Ubica uno de los agujeros de la cubierta de la caja y atornilla el perno (6). Recogiendo un segundo perono, repite la operación (7). Después de haber insertado los cuatro pernos, el robot saca la caja y la coloca sobre la mesa.

mas. Cada nivel de la jerarquía de control acepta órdenes del nivel inmediato superior y responde dando una serie de órdenes al nivel inmediato inferior, utilizando, en el caso que sea oportuno, una realimentación sensorial para cerrar el bucle de control. Supongamos, por ejemplo, que se trata de controlar un manipulador para un trabajo completo cual es el ensamblaje de un pequeño motor de explosión. El trabajo puede descomponerse en una serie de tareas más simples, tales como buscar las piezas *a* y *b*, encajar *a* en *b* anulando las fuerzas no axiales, sujetar la pieza *c* al conjunto *ab*, y así sucesivamente. Cada tarea de éstas puede, a su vez, descomponerse en una serie de movimientos elementales, tales como alcanzar, asir, seguir una trayectoria prefijada y otros parecidos. Estos mismos movimientos elementales se pueden dividir en una serie de posiciones del manipulador, expresadas en coordenadas x, y, z, y en las orientaciones de las pinzas. Finalmente, las posiciones del manipulador se pueden transformar en unas secuencias coordinadas de voltajes que gobiernen los servomecanismos accionadores de las articulaciones.

En general, hay dos tipos de señal de entrada para cada nivel de la jerarquía. En primer lugar, hay las órdenes que proceden de un nivel superior. En el nivel máximo, la señal de mando puede provenir de un operador humano o de algún



ROBOT UNIMATE, construido por Unimation, Inc., de Danbury, Connecticut, es capaz de ejecutar seis movimientos básicos, programables separadamente por un controlador manual. La pinza o "mano" del robot sujeta luces de colores diferentes para identificar a los seis movimientos. El brazo de pistón desarrolla tres movimientos: dentro-fuera *(azul)*, arriba-abajo *(blanco)* e izquierda-derecha *(rojo)*. Para cada posición del brazo del robot, las pinzas ejecutan otros tres movimientos: inclinación *(naranja)*, giro *(verde)* y deslizamiento *(amarillo)*.

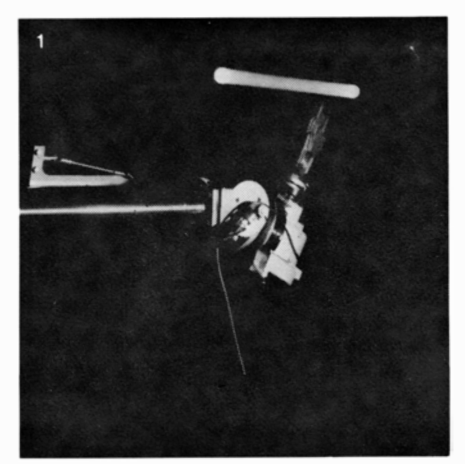
otro mecanismo externo que fije el objetivo. En todos los demás niveles, las señales de control son las salidas que provienen del nivel jerárquico inmediatamente superior.

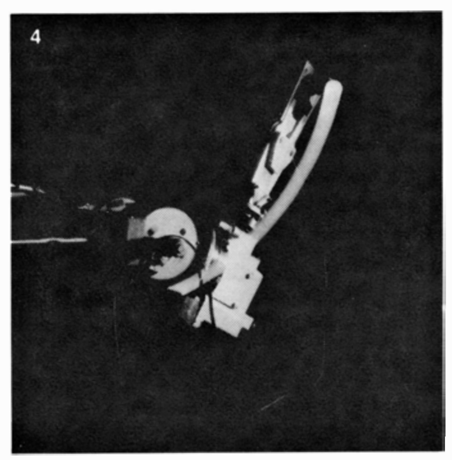
La segunda clase de entradas consiste en señales de realimentación que indican el estado del manipulador y del entorno. Tales señales pueden describir la posición y el movimiento de las articulaciones de fuerza, de contacto o de proximidad de las pinzas del robot. Además, los datos de la realimentación pueden suministrar señales de medida de tiempo procedentes de máquinas externas, con las que el robot tenga que interaccionar. Los datos de la realimentación pueden describir, también, la situación de los obstáculos, la forma y la posición de las piezas a manejar y el movimiento del sistema de referencia, o suministrar cualquier otro tipo de información relativa a la tarea de manipulación del robot. Sobre todo, en los niveles superiores de control, tales informaciones de entrada pueden requerir un complicado tratamiento de señales ópticas, acústicas, táctiles u otras, semejantes a las que perciben normalmente los sentidos humanos.

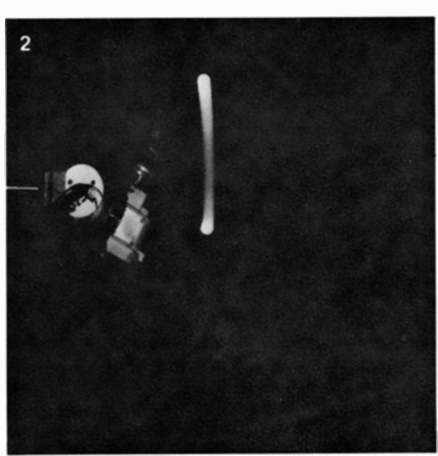
Cuando se considera desde una perspectiva jerárquica el control de un robot proyectado para la ejecución de una labor compleja, se requiere que cada nivel de control sea capaz de transformar las órdenes de un nivel superior en una secuencia de órdenes para el nivel inferior siguiente. La dificultad técnica para conseguirlo deriva, sobre todo, del volumen y complejidad de los datos de realimentación que deben manejarse en los cálculos. A su vez, el tipo de realimentación depende mucho del grado de indeterminación que presenta el entorno.

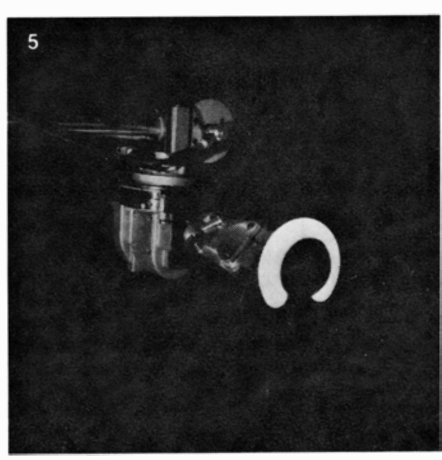
Si los desajustes se limitan a unos pocos milímetros en la posición y a unos pocos grados en la orientación, se puede dotar al robot de sensores de fuerza, de contacto o de proximidad que suministren la información correctora necesaria. Unas veces, las fuerzas se miden mediante la detección de variaciones de las corrientes de los motores o por sistemas hidráulicos de presión diferencial colocados en los actuadores apropiados. En otros diseños, los sensores de fuerza consisten en galgas extensométricas acopladas a las "muñecas" del robot. O, también, se pueden montar los sensores de fuerza en la mesa sobre la que se halla la pieza en la que el robot introduce espigas y cojinetes, monta juntas y aprieta tornillos. Las fuerzas transmitidas a la mesa a través de la pieza que se está tratando se pueden usar para controlar tales operaciones de ensamblaje. La mayoría de los sensores táctiles son, simplemente, unos microinterruptores que transmiten una señal cuando dos superficies entran en contacto.

Los sensores de proximidad suelen ser más complicados. Los sistemas típicos emplean un rayo láser o un haz de radiación infrarroja y determinan la proximidad midiendo la energía reflejada por el objeto. Las distancias superiores a un metro pueden calcularse midiendo el tiempo de tránsito de un impulso de láser. Un robot que esté equipado con sensores de proximidad bien diseñados puede buscar objetos cuya posición no sea segura, y moverse hacia ellos a gran velocidad sin peligro de que ocurra ninguna colisión. Puesto que los sensores de fuerza, de contacto y de proximidad suministran información en un sistema de coordenadas solidario con el dispositivo sensorial, la información deberá entrar en la jerarquía de control a un nivel igual

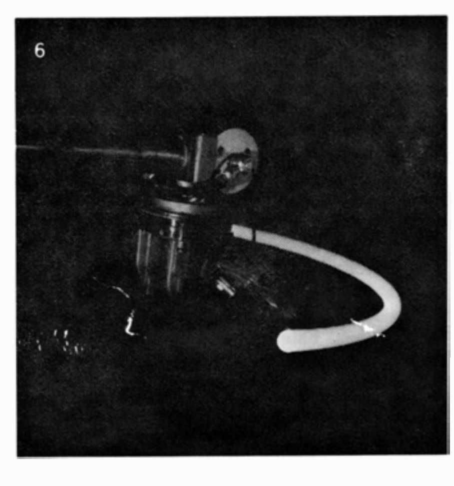










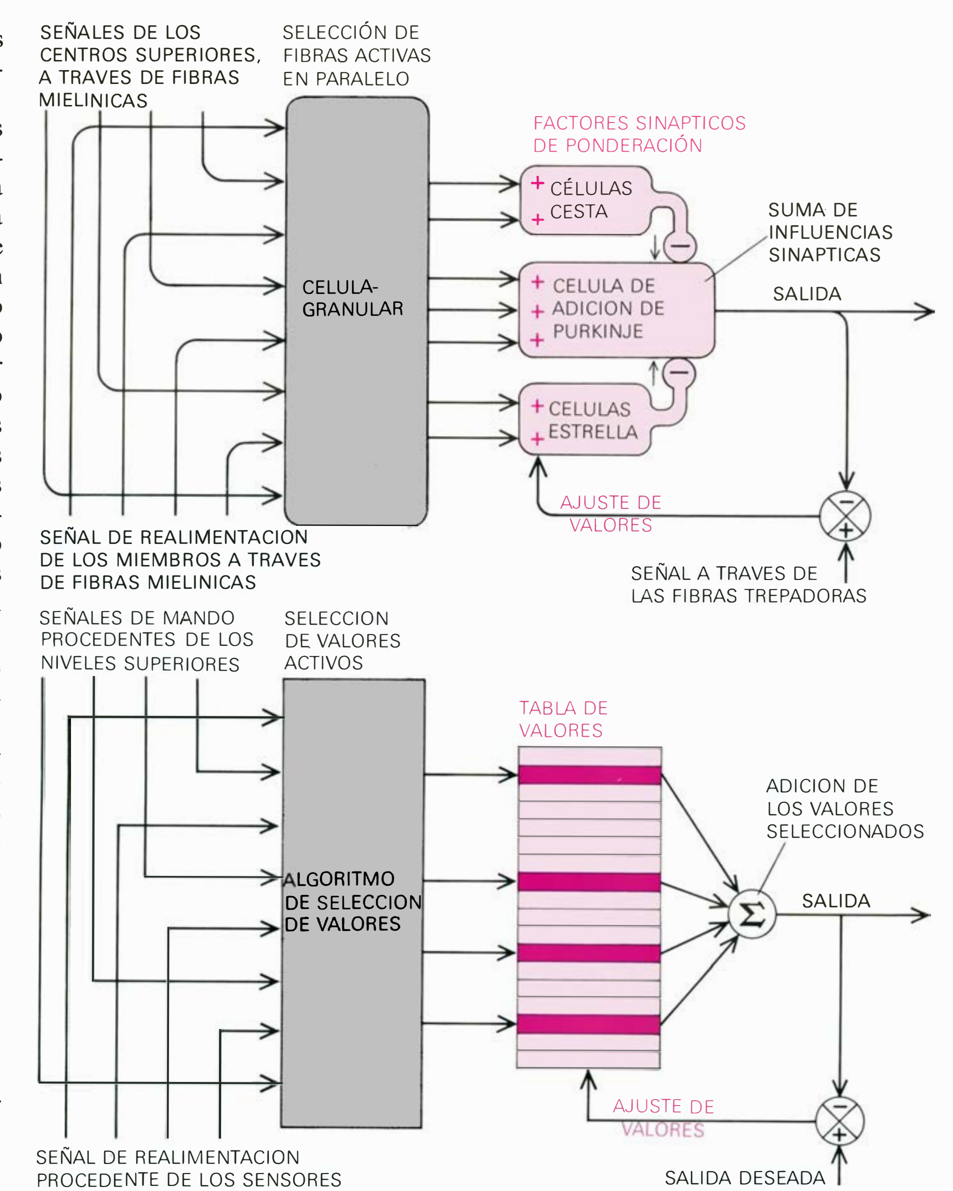


SEIS MOVIMIENTOS DEL ROBOT UNIMATE mostrados por separado. En la columna de la izquierda se muestran los tres movimientos básicos de los brazos: dentro-fuera *(1)*, arriba-abajo *(2)* e izquierda-derecha *(3)*. La columna de la derecha ofrece los tres movimientos de las pinzas: inclinación *(4)*, giro *(5)* y deslizamiento *(6)*. Del conjunto de las seis fotografías de esta ilustración, solamente la número cinco muestra el recorrido total del movimiento que ha efectuado el robot.

o superior a aquel en el que las señales de control puedan expresarse en coordenadas de sensor.

Existen varios métodos mediante los cuales se pueden utilizar los datos de realimentación, a diferentes niveles de la jerarquía de control, para compensar la indeterminación del entorno. Uno de ellos estriba en determinar la posición exacta de un objeto, y trasladar, girar o extender las trayectorias que habían sido registradas previamente para compensar los errores de posición medidos. Otro método consiste en introducir los datos de la realimentación en las ecuaciones matemáticas, usadas para calcular las trayectorias de la posición final del manipulador. Esta segunda técnica ha sido empleada por Whitney y James L. Nevins en el Draper Laboratory para obtener un manipulador, realimentado por fuerzas, cuyo objetivo es reconocer superficies, seguir perfiles e introducir espigas en orificios de tolerancia muy pequeña. Pueden utilizarse también señales de realimentación para realizar misiones secundarias condicionales en los programas de control. Esta posibilidad se ha demostrado en el Artificial Intelligence Laboratory, de la Universidad de Stanford en el montaje de una bomba de agua por parte de un robot.

Otra técnica para incorporar la realimentación a la jerarquía de control del "controlador de articulación de modelo cerebelar" (CMAC), por uno de los autores (Albus) en el National Bureau of Standards. El CMAC es un sistema adaptativo que utiliza un sistema de almacenamiento distribuido para memorizar o aprender la salida correcta que corresponde a cada combinación adecuada de señales de entrada. Al igual que el cerebelo, en el que se ha inspirado el modelo, el CMAC puede hacer frente a un gran número de variables de mando y de realimentación y presenta muchas características del reflejo condicionado. Resulta de gran utilidad potencial a distintos niveles en la jerarquía de control.



CONTROLADOR DE ARTICULACION SEGUN MODELO CEREBELAR (CMAC). El controlador de articulación trata de imitar el esquema de tratamiento de la información que sigue el cerebelo, esquema que aparece en la ilustración superior. David Marr, actualmente en el Instituto de Tecnología de Massachusetts, fue quien comenzó a desarrollar el modelo cerebelar, elaborado más tarde por uno de los autores (Albus). Las señales de entrada, procedentes de los centros superiores del cerebro, y las señales de realimentación procedentes de los distintos miembros, se transportan por fibras mielínicas hasta el estrato de células granulares, que actúan como de codificador de direcciones de posición. Cada señal de entrada constituye una dirección que sólo es decodificada por un conjunto determinado de fibras, conectadas en paralelo. Los impulsos, transmitidos por las fibras paralelas, viajan a través de varios caminos sinápticos y dan lugar a la señales de salida, las cuales emergen de las células de Purkinje. Cada salida puede considerarse como el contenido de una dirección de las fibras mielínicas. El valor que tenga este contenido depende de la magnitud de los factores sinápticos de ponderación *(indicados con el signo aritmético más)*, los cuales son ajustables. La ponderación actúa de tal modo que adapta los datos aportados por las fibras ascendentes a los de salida de las células de Purkinje. De esta forma el cerebelo puede mantener una relación funcional entre entrada y salida. En el diagrama inferior se ha representado la versión del CMAC realizada con ordenador. Cada entrada selecciona un solo conjunto de valores, los cuales se suman para obtener una salida. El proceso de selección actúa de forma tal que si hay dos entradas parecidas, resultarán seleccionados conjuntos de valores parcialmente solapados. De este modo, siempre que los valores seleccionados por una dirección de entrada resulten ponderados para modificar la salida, también cambiarán algunas de las salidas vecinas a causa del solapamiento de los valores seleccionados. Este solapamiento permite que el CMAC haga generalizaciones entre entradas similares a cómo los seres humanos y otros animales generalizan a partir de experiencias semejantes entre sí. La siguiente ilustración muestra algunos ejemplos de operación del sistema CMAC.

Cuanto mayores sean las indeterminaciones en el mundo del robot, tanto más complejas tendrán que ser las técnicas de realimentación y de control. Por ejemplo, si el robot ha de manipular objetos que estén colocados al azar o apilados en un montón, una simple información de fuerza, de contacto y de proximidad resultará a buen seguro insuficiente. En algunos casos extremos, se requerirá que remueva un obstáculo al menos para alcanzar el objeto que está buscando. Y puede requerirse incluso que haya de seleccionar las herramientas con las cuales conseguir sus objetivos. Tales maniobras exigen un robot dotado de cierta capacidad de planificación racional.

Durante muchos años, los investigadores en el campo de la inteligencia artificial se han esforzado por equipar al robot de unas facultades intelectuales modestas. En la mayor parte de los casos este empeño ha exigido la incorporación, en los niveles superiores de control, de una cierta estructura de información interna –un "modelo del mundo" o conjunto de conocimientos– con el fin de poder representar, de un modo significa-

tivo, el estado de su entorno. El robot debe estar dotado asimismo de un adecuado repertorio de sensores y de circuitos procesadores de datos, para analizar el medio que le rodea y mantener así constantemente actualizado su conjunto interno de conocimientos.

Cuando el robot se encuentra con que debe ejecutar una orden de entrada, intenta establecer un modelo hipotético del mundo deseado. A continuación trata de planear un conjunto de actuaciones que permitan convertir el modelo del mundo existente en el modelo deseado. Cada fase de estas actuaciones internas debe corresponder a órdenes de salida que puedan ejecutarse por un nivel inferior de la jerarquía. Una vez encontrado tal proceso interno, puede ser convertido en la serie necesaria de órdenes de salida.

Esta técnica se utilizó en el robot llamado Shakey, construido en el Stanford Research Institute. A Shakey se le podían encomendar trabajos tales como encontrar una caja de tamaño, forma y color determinados, y trasladarla a una posición prefijada. El robot era capaz de buscar la caja en varias salas, sortear diversos obstáculos y planear una forma de actuación adecuada al caso. En algunas situaciones, se requería llevar a cabo una acción preliminar antes de que pudiera alcanzarse el objetivo principal. Una vez, Shakey descubrió que, si movía una rampa unos cuantos decímetros, podía subirse a la plataforma en la que se había colocado la caja. Quizás sea superfluo apostillar que, cuando se quiere dotar a un robot de los medios necesarios para hacer frente a semejantes tareas, son muchos los problemas a resolver.

Un modo prometedor de afrontar el problema del análisis automático de las imágenes y las escenas es el desarrollado en el Laboratory for Computer Science, del M.I.T. Los datos de un analizador de televisión se elaboran por unos programas de ordenadores que reconocen perfiles, definen superficies, analizan formas y componen las estructuras de los datos para especificar cómo encajan entre sí todas esas cualidades elementales, para dar una descripción significativa de los objetos reales del entorno que se considera. En los trabajos más recientes realizados en el Artificial Intelligence Laboratory del M.I.T., los programas de análisis de escenas se han organizado en una estructura no jerárquica, de forma que todo subprograma pueda llamar a cualquier otro e intercambiar información con él. Así, por ejemplo, un subprograma de reconocimiento de perfiles

puede llegar a una conclusión provisional y comunicar directamente su resultado al programa que le había solicitado, o bien decidir que necesita la asistencia de otros subprogramas. El diálogo comprende todo aquello que se podrían denominar consejos, sugerencias, observaciones, reclamaciones, críticas, respuestas, errores y conjeturas.

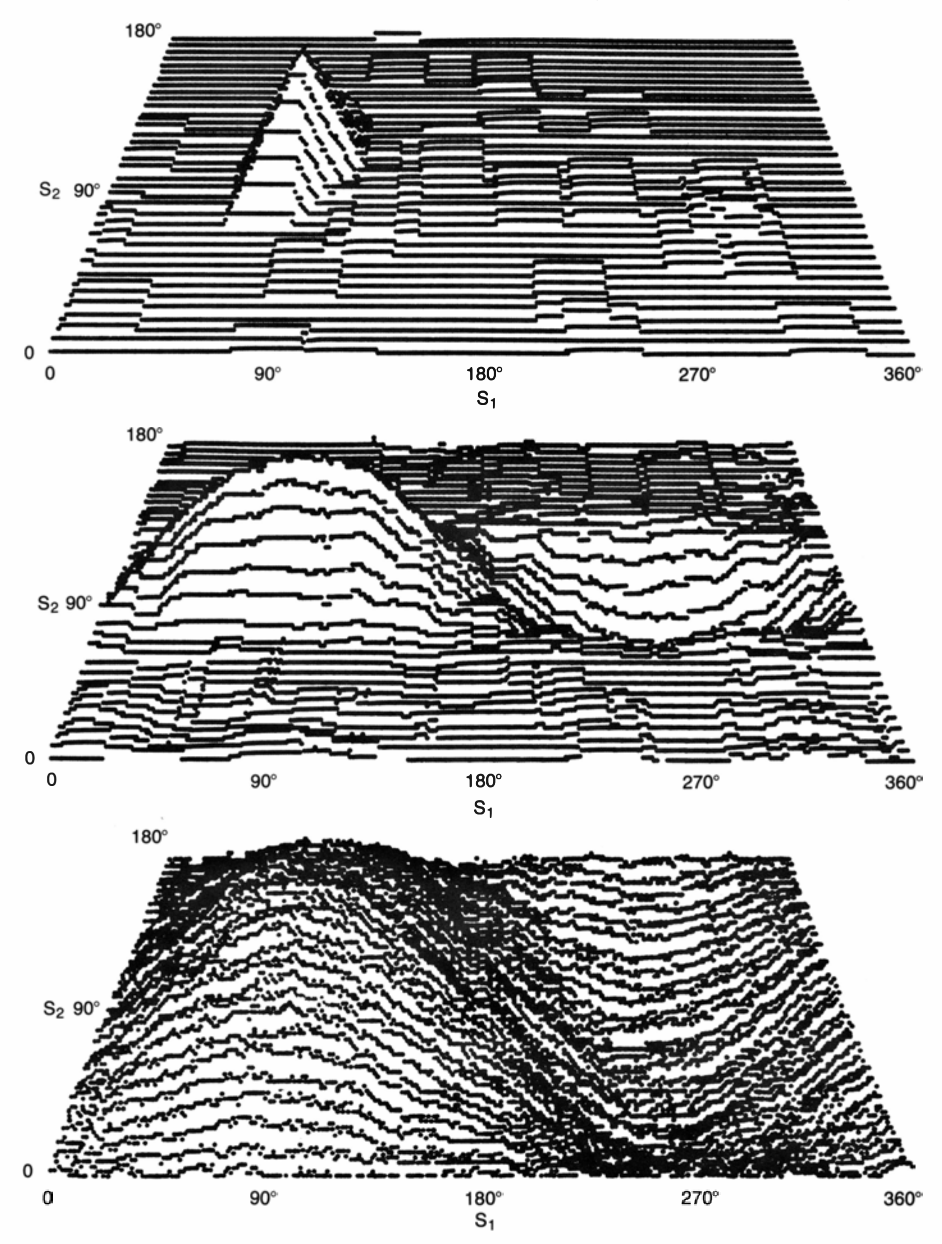
A pesar de que tales investigaciones han proporcionado una información muy valiosa sobre complicados procesos de percepción visual, no parece probable que los programas analíticos específicos desarrollados hasta hoy puedan incorporarse a los sistemas de robot de tipo práctico hasta dentro de muchos años. Los programas no sólo necesitan un gran poder de cálculo, sino que requieren también que los objetos existentes en el entorno tengan formas geométricas sencillas y resalten del fondo.

En el Stanford Artificial Intelligence Laboratory, Thomas O. Binford y colaboradores han adoptado un enfoque completamente distinto respecto al problema de la capacidad visual de las máquinas. Iluminan los objetos del entorno con un rayo láser muy fino al que una lente cilíndrica hace dispersarse en un plano. Dondequiera que el plano luminoso corte la superficie de un objeto, el contorno de éste queda destacado merced a un fino perfil luminoso. Moviendo el plano en diferentes posiciones y orientaciones se puede ir construyendo un modelo compuesto que indica la forma tridimensional de la superficie visible. Luego, el modelo se analiza por un programa de ordenador que intenta deducir cuál es la estructura del objeto entero a partir de las superficies visibles para el sistema de exploración. Este método presenta la ventaja de que evita muchas complicaciones causadas por las sombras y las formas irregulares.

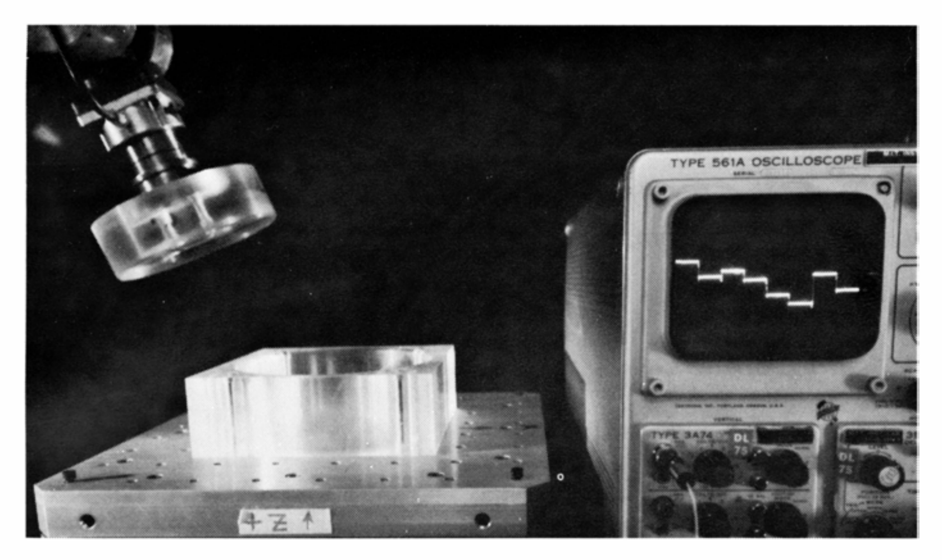
En el Jet Propulsion Laboratory, del Instituto de Tecnología de California, se está desarrollando un sistema de robot con capacidad visual para localizar y recoger rocas de una superficie que simula ser la de Marte. El sistema utiliza una cámara de televisión y un programa de análisis de escenas para detectar el contorno de una roca. Luego se utiliza un telémetro de láser para medir la posición exacta de la roca, para que el manipulador pueda alcanzarla.

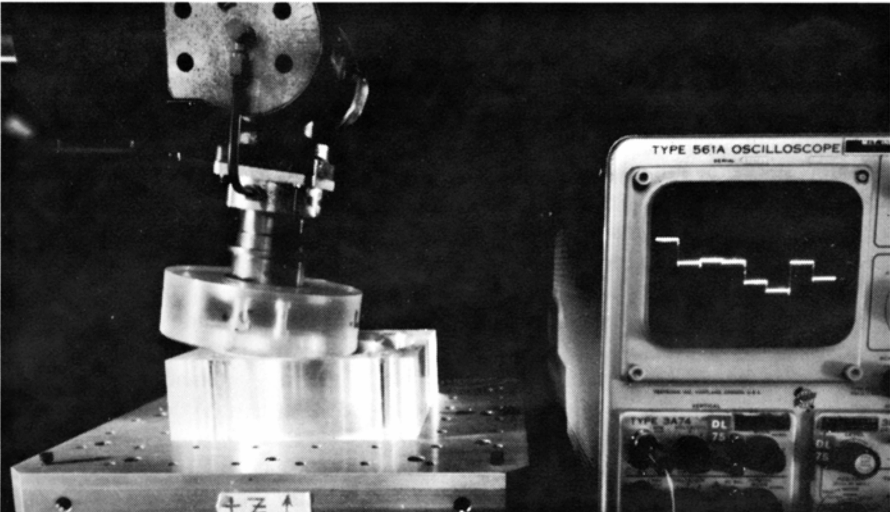
En el Stanford Research Institute, Charles A. Rosen y David Nitzan, entre otros, empeñados en descubrir procedimientos de análisis visual que resulten adecuados para su aplicación industrial, han ideado un sistema de exploración que consiste en una línea de fotodiodos montados sobre una cinta transportadora. A medida que los objetos (que están orientados al azar y no son necesariamente iguales entre sí) van pasando en fila de uno en uno, los fotodiodos construyen la imagen del contorno de cada uno de ellos. Unos programas de análisis visual, bastante sencillos, determinan el contorno del objeto, miden su posición y su orientación y llevan a cabo una inspección elemental de posibles defectos. Tanto si el objeto vale como si debe ser rechazado, se dan las oportunas órdenes al manipulador, el cual toma la pieza móvil y la deja en una orientación adecuada para ulteriores tratamientos.

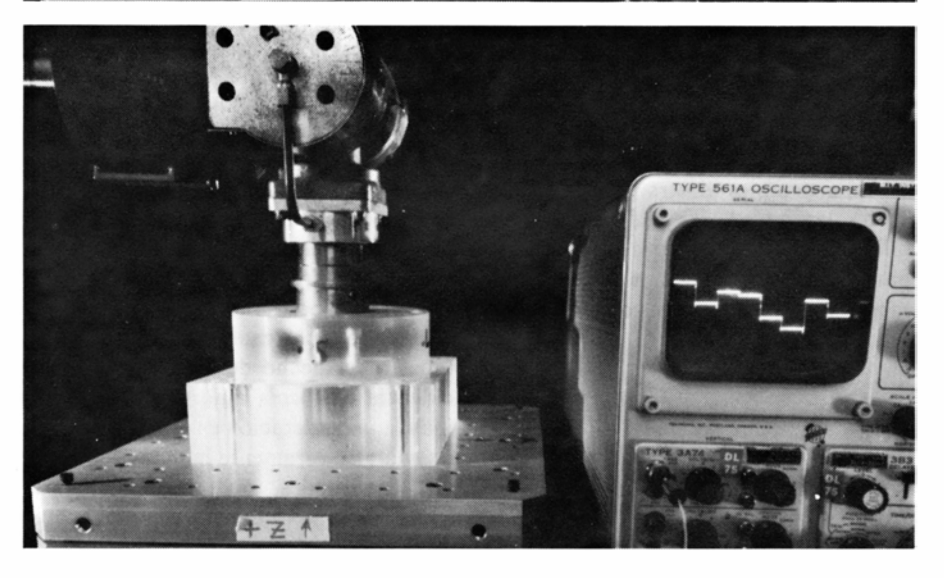
Habrán de transcurrir indudablemente varias décadas antes de que los robots puedan simular los procesos básicos del comportamiento inteligente, y hasta realizar la elaboración de la información sensorial requerida para la captación visual. Pasará también algún tiempo an-



EL CONTENIDO CAMBIANTE DE LA MEMORIA CMAC aparece en tres figuras impresas por un ordenador. En ellas se observa cómo la memoria del CMAC generaliza entre entradas parecidas. El pictograma superior muestra cómo se modificó la memoria mediante una sola entrada, +1, aplicada en la dirección de entrada $S_1=90$, $S_2=90$. Esta entrada altera el contenido de la memoria no sólo en la dirección seleccionada sino también en las direcciones vecinas. Como puede observarse, el efecto sobre las direcciones vecinas va disminuyendo proporcionalmente a la distancia entre éstas y la dirección seleccionada. El pictograma central muestra el resultado de introducir una onda sinusoidal formada solamente por 16 entradas dispuestas a lo largo del eje $S_2=90$. También aquí, las direcciones adyacentes al eje quedan alteradas a consecuencia del proceso de generalización. El pictograma inferior muestra el contenido de la memoria del CMAC cuando la función es producto de dos ondas sinusoidales, una a lo largo de cada eje. Gracias a la generalización, sólo fue necesario seleccionar 175 entradas para almacenar esta función en un espacio de 360×180, que comprendía por tanto 64 800 posiciones. Las ilustraciones aquí expuestas muestran de qué modo una función continua acotada puede almacenarse en la memoria y, por lo tanto, emplearse como salida del CMAC, mediante la introducción de datos en un número de puntos relativamente pequeño.







PLATAFORMA SENSORA DE FUERZA, que puede usarse para dirigir a un robot en una tarea de montaje, como la introducción de un cilindro en una abertura ajustada. La técnica se ha desarrollado en el Charles Stark Draper Laboratory, en donde se han obtenido estas fotografías. El grabado superior muestra al robot acometiendo su primer intento en la tarea de montaje. La imagen escalonada que aparece en la pantalla del osciloscopio representa las fuerzas que actúan sobre ocho galgas extensométricas que hay colocadas encima de la mesa y que se encuentran debajo del bloque cuadrado. Cada escalón, en la imagen, corresponde a la fuerza que actúa sobre un sensor. Los sensores responden tanto a las fuerzas como a los pares. Cuando el robot trata de encajar el cilindro en el bloque y falla ***(centro),*** **los sensores detectan con exactitud dónde y cómo presiona el cilindro contra los lados del agujero y la imagen del osciloscopio es distinta a la que se produce en reposo (ausencia de fuerzas). Un programa de ordenador analiza el desajuste y calcula la trayectoria que deberá seguir el brazo del robot para que el cilindro entre suavemente en el agujero la vez siguiente** ***(abajo).***

tes de que los robots sean capaces de inventar estrategias con las que enfrentarse ante nuevas situaciones con posibilidades parecidas a las de un niño de dos años. La captación visual y la invención de estrategias son procesos extremadamente complejos que no podrán comprenderse hasta que no se conozca mucho mejor el funcionamiento fundamental del cerebro humano.

Una cuestión muy distinta, sin embargo, son las posibilidades del robot para acometer trabajos útiles en el medio altamente organizado de una fábrica moderna. En una fábrica no se exige que el robot disponga de facultad creadora, ni que tenga un grado de discernimiento significativo. Los criterios básicos a tener en cuenta en un robot industrial son los siguientes: que sea barato, fiable y fácilmente programable para que pueda ejecutar una serie de operaciones muy concretas. Existen métodos prácticos ordenados a que el robot pueda apreciar y compensar las variaciones de las medidas normales y los pequeños desajustes de las piezas. En el entorno restringido de una fábrica industrial, las órdenes de entrada del nivel máximo pueden estereotiparse completamente, al igual que lo están las comunicaciones que hay entre el ordenador y los dispositivos periféricos convencionales, tales como las memorias de disco. Por tanto, podrían generarse series de señales de control de alto nivel mediante compiladores y otras técnicas de programación harto conocidas. Ya se han desarrollado varios lenguajes de programación con el fin de programar y manejar máquinas herramientas controladas numéricamente. Recientemente varios grupos de ingenieros han empezado a crear lenguajes parecidos para los sistemas de robot.

Los progresos descritos aquí sugieren que no existen barreras técnicas fundamentales que impidan conseguir fábricas altamente automatizadas en un futuro próximo. En Japón, el gobierno, la industria y las universidades trabajan en un importante programa para el desarrollo de una planta prototipo, sin personal, que utilizaría la tecnología de los robots para fabricar los componentes de las máquinas herramientas. Se prevé que la planta entrará en funcionamiento hacia el año 1980 y, si tiene éxito, servirá de modelo para otros proyectos industriales. Por lo tanto, parece muy probable que dentro de unos quince años estén en servicio, por lo menos en algunos países, fábricas altamente automatizadas mediante la utilización de robots, y que quizá tengan, incluso, una capacidad parcial de autorreproducción.

El problema del cáncer

Parece ser que casi todos los tumores malignos están causados por factores ambientales. Por consiguiente, la vía de investigación más eficaz será la que conduzca a la detección y eliminación de esos factores carcinógenos

John Cairns

Durante los últimos 150 años, el mundo occidental ha ido eliminando prácticamente las enfermedades infecciosas como una de las causas de mortandad más significativas. Un niño nacido hoy en los Estados Unidos puede esperar tener una vida que no se vea perturbada por enfermedades tales como la escarlatina, la difteria, la tuberculosis, la fiebre tifoidea y la disentería, que eran las principales causas de muerte en las tres o cuatro últimas generaciones. La esperanza de vida ha tenido un incremento constante desde mediados del siglo diecinueve. Y esa prolongación del promedio de vida débese, por encima de todo, al mejoramiento de los servicios de higiene y sanidad; los frutos más espectaculares de la investigación científica, como la introducción de vacunas y antibióticos, se han limitado a completar el proceso.

La incidencia de muerte tiene ahora su cota superior en la vejez, y puede atribuirse, por tanto, a enfermedades que son peculiares de la edad avanzada o letales solamente para los ancianos. Aunque muchísimos cambios de los que ocurren en el cuerpo con el avance de la edad podrían clasificarse como enfermedades, hay dos condiciones particulares que suelen aparecer en la edad avanzada, y que son con frecuencia causa directa de muerte: la enfermedad arterial (ateroma y arteriosclerosis) y el cáncer. La enfermedad arterial es letal cuando afecta a las arterias que irrigan el corazón o el cerebro; tiene en su haber actual el cincuenta por ciento, aproximadamente, de la mortalidad absoluta norteamericana. El cáncer es letal cuando se disemina a partir de su lugar de origen: suma casi el veinte por ciento de la misma muestra.

La gente se comporta de manera muy distinta ante un tipo u otro de estos dos grupos de enfermedades. No sólo porque la muerte por enfermedad arterial suele tener un curso rápido mientras que la muerte por cáncer puede prolongarse dolorosamente. Por alguna razón, el ataque al corazón y al cerebro tienden a concebirse como riesgos propios de la edad avanzada y el fin normal de una vida satisfactoriamente larga, o bien, cuando ocurren en hombres de mediana edad, como el precio de la intemperancia en la comida y la falta de ejercicio. Por el contrario, se considera el cáncer como una enfermedad imprevisible que ataca, indiscriminadamente, a ricos y pobres, obesos y delgados, ancianos y hombres en plenitud de facultades como si no guardara relación con causa externa alguna. Si todo ello fuese cierto, nuestra única esperanza de dominar el cáncer sería perfeccionar el tratamiento de la enfermedad. Uno de los propósitos de este artículo es mostrar que la mayoría de las neoplasias más extendidas hallan su causa en factores ambientales; como quiera que podemos modificar el ambiente, esos cánceres son potencialmente evitables.

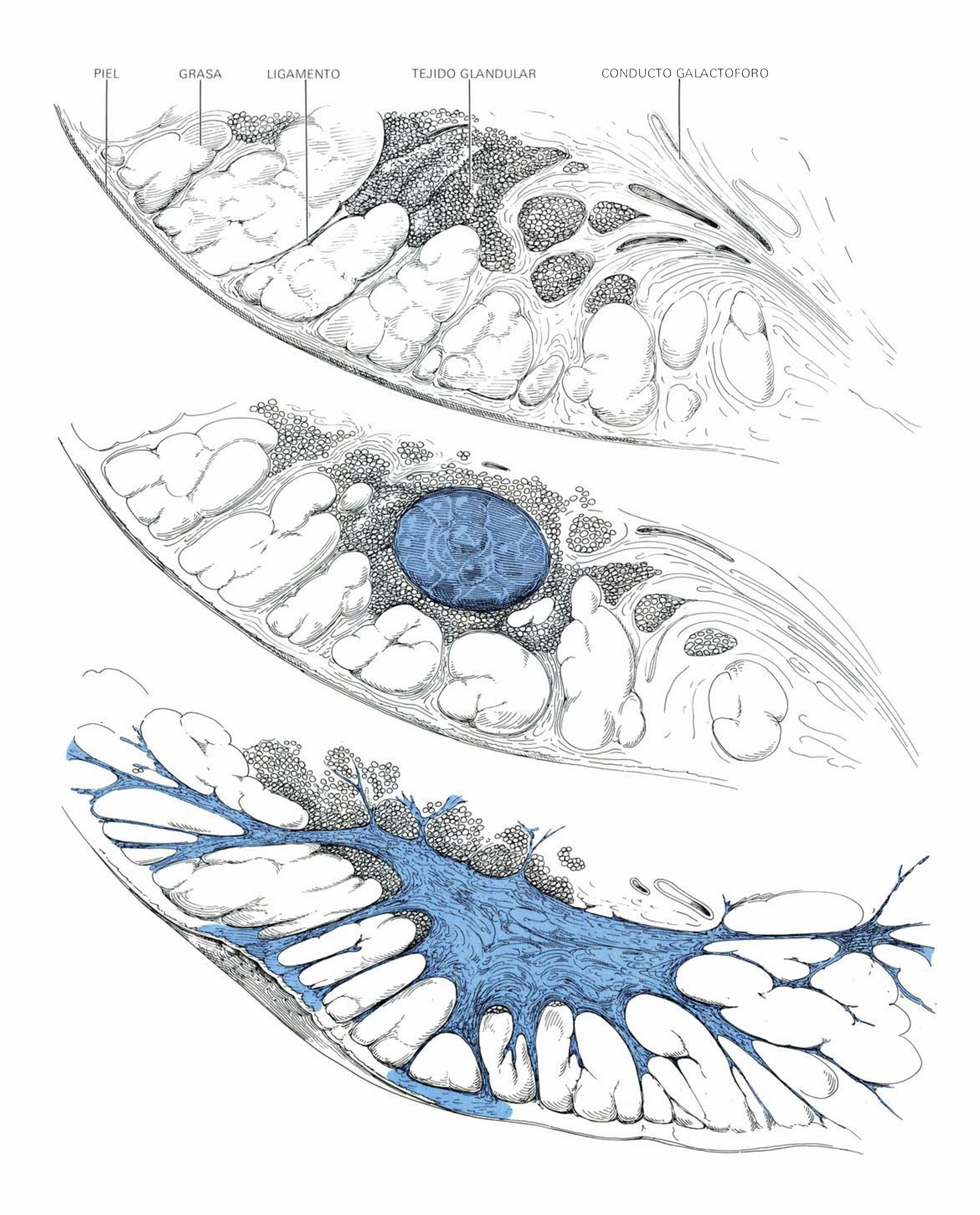
Incidencia

En cualquier parte del cuerpo pueden originarse grupos de células que proliferen de forma anormal. De ellos, los que no pueden invadir los tejidos vecinos y permanecen estrictamente localizados reciben el nombre de tumores benignos; los que se diseminan desde el lugar de origen y pueden alcanzar el torrente sanguíneo y el sistema linfático reciben el nombre de tumores malignos, o cánceres.

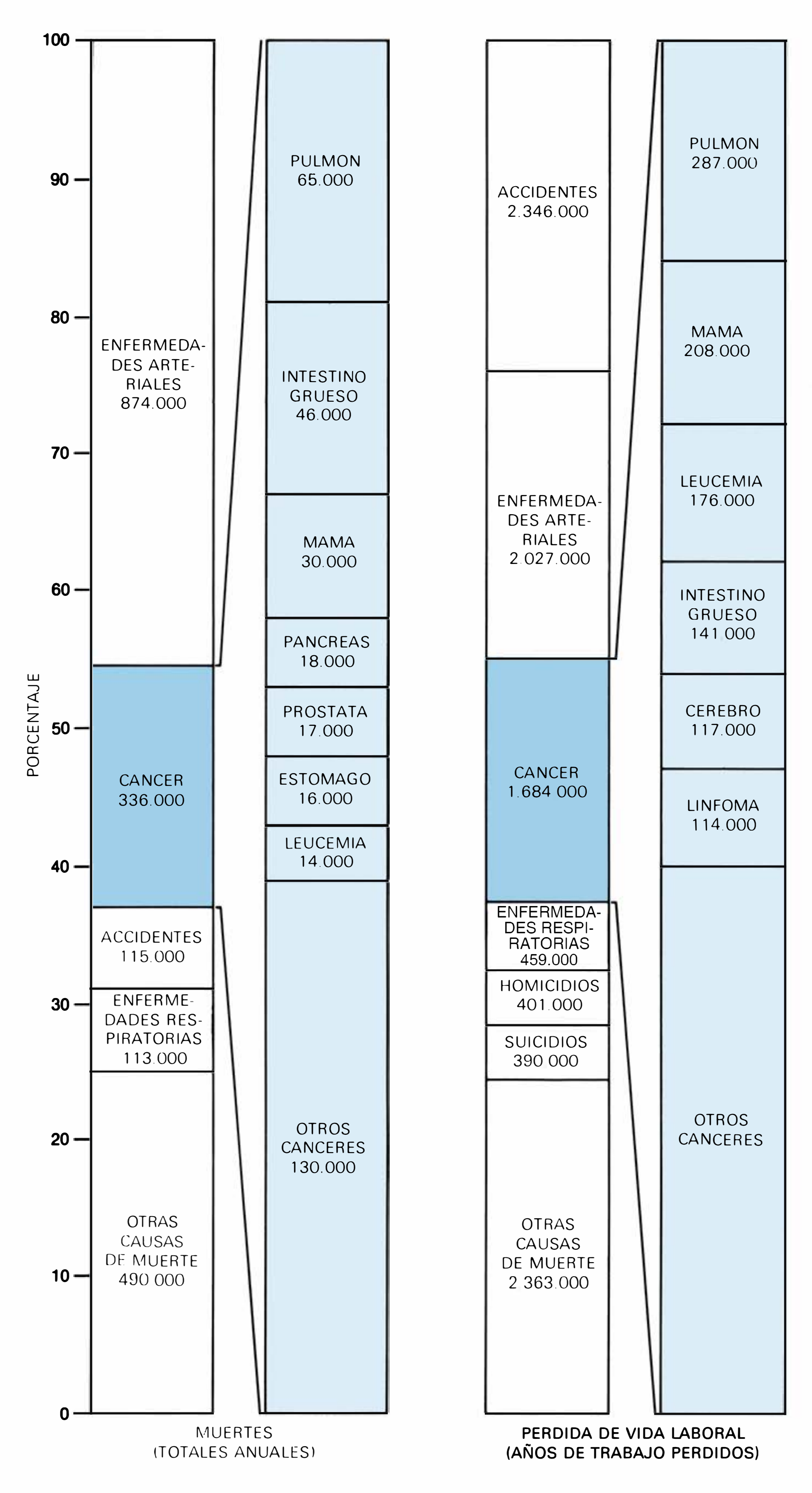
Los tipos de cáncer se encuadran en tres extensos grupos: carcinomas, sarcomas y leucemias. Los carcinomas se originan en los epitelios, capas de células que recubren la superficie del cuerpo y tapizan las diferentes glándulas. Los sarcomas, mucho más raros, se originan en las estructuras de sostén, como el tejido fibroso y los vasos sanguíneos. Las leucemias y los linfomas se originan en las células hematopoyéticas de la médula ósea y de los linfonodos. Los términos –carcinoma, sarcoma y leucemia– se hallan tan arraigados en el lenguaje común que se hace obligada su mención, pero mi intención es sobreentender con su uso la de tres formas de carcinogénesis fundamentalmente distintas, las cuales postulan métodos distintos para su prevención y tratamiento. Esa clase de información sólo puede obtenerse a través de un sistema de clasificación más preciso.

Los distintos tipos de cáncer se clasifican principalmente por el órgano en que se han originado y por la clase de células interesadas. Así considerada, se cuenta hasta un centenar de variedades de la enfermedad. Tan elaborada clasificación carecería de interés general si no fuese porque a cada variedad corresponde una causa diferente, como muestra el hecho de que cambie su incidencia de forma autónoma cuando se induce una alteración en el ambiente. Por presentarse raramente la mayoría de esas cien variedades, podremos abarcar la gama de causas de muerte por cáncer ciñéndonos a una serie de enfermedades bastante restringida.

La mitad aproximadamente de todas las muertes por cáncer vienen provocadas por tumores malignos localizados en tres órganos: el pulmón, el intestino grueso y la mama *(véase la ilustración de la pág. 90)*. No habrá, consiguientemente, una reducción significativa en la gráfica de la mortalidad total por cáncer mientras no se encuentre algún medio para curar o prevenir estos tres



EL CRECIMIENTO DE UN TUMOR en la mama amenaza generalmente la vida sólo cuando el tumor puede diseminarse y alcanzar partes distantes del cuerpo. La mama normal *(arriba)* consta de tejido glandular, adiposo y de otras estructuras. Los tumores se desarrollan casi exclusivamente en el tejido glandular; están compuestos por células en las que se han eliminado los controles normales del crecimiento y de la proliferación. Un tumor benigno *(en el centro)* puede crecer rápidamente y alcanzar un tamaño considerable, pero no puede salirse del tejido en que se desarrolla. Un tumor canceroso *(abajo)* puede infiltrarse por todo el tejido glandular, afectar frecuentemente a los ligamentos y la piel y, a veces, penetrar en el músculo que subyace a la mama. Además, el cáncer puede emigrar en algunos casos a través de la sangre o del sistema linfático para establecer nuevas colonias de células en órganos distantes y extraños. A este proceso se le da el nombre de metástasis (que significa paso de un lugar a otro). La responsable de la letalidad de los tumores malignos es su dispersión metastásica.



LA MORTALIDAD POR CANCER constituye hoy, aproximadamente, la quinta parte de todos los casos de muerte en Estados Unidos. Entre las muertes atribuidas al cáncer, más del sesenta por ciento son causadas por las neoplasias más frecuentes. La diferencia porcentual restante se reparte entre más de cien formas raras de tumores malignos. La incidencia de cada tipo de cáncer puede evaluarse también atendiendo a la reducción de vida que causa, que se mide aquí en años de trabajo suponiendo que la vida laboral se extienda desde los veinte hasta los sesenta y cinco años. El principal efecto de la consideración de la reducción de vida en vez de la simple mortalidad es aumentar la importancia de los accidentes del total de muertes; entre los distintos tipos de cáncer, destaca la elevada incidencia relativa de leucemias en la juventud, aunque predomina el cáncer de pulmón.

tipos de cáncer. Cada uno de ellos puede ser considerado como una entidad discreta (sin relación de continuidad con los otros dos), porque la frecuencia de uno varía con el cambio de los factores ambientales, sin que por ello tengan que variar los demás.

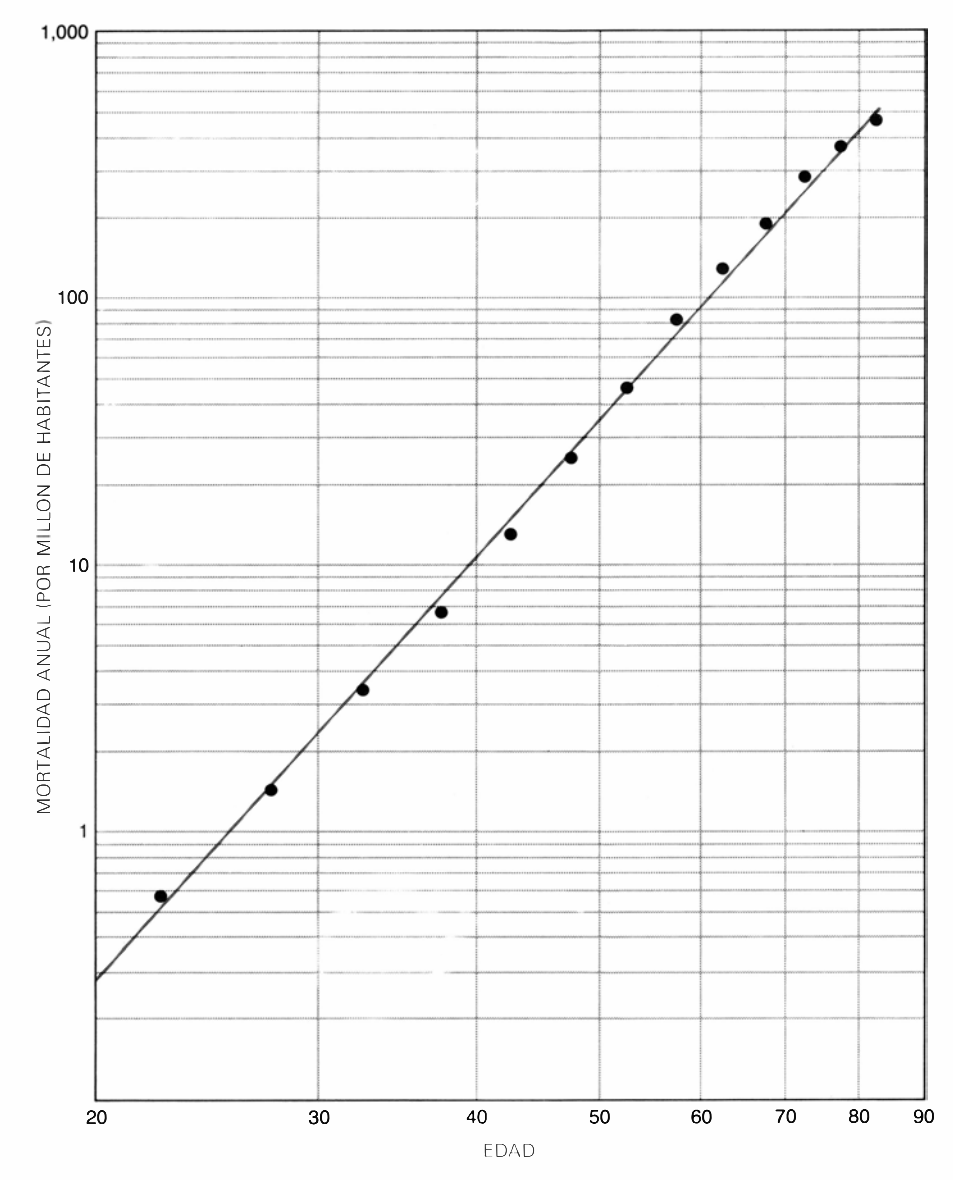
Podríamos afirmar, con razón, que no importa tanto el número absoluto de muertes cuanto la pérdida en cantidad de vida. La muerte por cáncer de próstata de un hombre nonagenario es menos trágica que la muerte por leucemia de un joven. A la hora de determinar prioridades, deberíamos considerar, por tanto, la distribución por edades de las víctimas de cada causa de muerte. Existen diversas maneras de hacerlo. Por ejemplo, se puede calcular en cuánto reduce la esperanza de vida media cada causa principal de muerte, o cada clase de cáncer; o bien, en concreto, en cuánto reduce cada una la vida laboral hasta la edad de 65 años. El resultado más destacado de tal sistema de estimación estriba en potenciar la importancia relativa de los accidentes, entre las causas generales de muerte, y de las leucemias y linfomas, entre los distintos tipos de cáncer. A pesar de todo, el cáncer de pulmón sigue encabezando la lista.

Todas estas estadísticas se refieren a la mortalidad, que constituye el patrón más accesible y seguro para medir la incidencia del cáncer. Lo que resulta tarea fácil es determinar la frecuencia de los distintos tipos de cáncer. En primer lugar hay que considerar aquellos pacientes que se presentan por sí mismo al médico y en el reconocimiento se les descubre dicha enfermedad que, en la mayoría de los casos, más pronto o más tarde, les causará la muerte. Esta aislada y deprimente estadística confirma la opinión de que la ciencia se halla desvalida ante enfermedades que ocurren principalmente en la madurez y en la ancianidad. El valor real de la incidencia sería más exacto si pudieran considerarse también las formas que pueden descubrirse a través de una investigación preventiva. Se han descubierto así muchas clases de cáncer en exámenes de rutina, poniéndose de manifiesto que pueden detectarse tipos de cáncer mucho menores y asintomáticos de lo que cabría esperarse. De ello se deduce que la mayoría de esos conglomerados de células invasoras tienen que regresar a su centro de origen antes de alcanzar un tamaño apreciable, o bien, y ello es más probable, crecen con tanta lentitud que no provocan síntomas durante la vida del paciente. Por ejemplo, la esti-

mación convencional de la incidencia del cáncer de próstata en varones de 70 años es, aproximadamente, 200 casos por 100.000 varones por un año, es decir el 0,2 por 100 anual; sin embargo, las autopsias rutinarias de hombres de 70 años fallecidos por otra causa han mostrado la existencia de microscópicos cánceres invasivos de la próstata con una frecuencia que oscila entre el quince y el veinte por ciento de los mismos. La incidencia medida por este método resulta así cien veces mayor.

Cuando la localización del cáncer permite el examen directo, los reconocimientos son mucho más fáciles de realizar y dan el mismo resultado. Por ejemplo, una reciente inspección en un distrito rural del estado de Tennessee mostró que un cuatro por ciento, aproximadamente, de la población adulta tenía cáncer de piel. Es más, parece probable que si pudiéramos examinar con la misma facilidad todo el organismo, encontraríamos que hacia el ecuador de su vida, el individuo ha ido adquiriendo varios nódulos de células proliferantes e invasoras que podrían definirse verosímilmente como cancerosos. Pero si no disponemos de otros muchos elementos de juicio en torno a la evolución natural de la enfermedad, es imposible prever qué tumores se diseminarán y cuáles no. Antes de discutir la evolución natural del cáncer, procede considerar la epidemiología de las formas más extendidas a fin de recabar datos interesantes sobre sus causas.

El primer paso para encontrar la causa de cualquier variedad particular de cáncer es determinar qué grupos de población muestran la incidencia más alta y qué los distingue del resto. Evidentemente, los ancianos son los individuos más afectados. Casi todas las clases de cáncer se dan con mayor frecuencia en la vejez, y la incidencia aumenta en relación directa con la edad. Como ejemplo típico la tasa de muerte por cáncer del intestino grueso aumenta unas mil veces entre los veinte y los ochenta años, registrándose gran parte del incremento a partir de los sesenta *(véase la ilustración de la derecha).*



LOS ANCIANOS componen la subpoblación que se halla más expuesta al riesgo del cáncer. La incidencia de casi todas las formas cancerosas crece dramáticamente con la edad. Aquí se expresa gráficamente la variación, por edad, de la tasa de mortalidad en Estados Unidos de un tipo de un tumor muy representativo, el de intestino grueso. Puede apreciarse que el logaritmo de la tasa de mortalidad se halla en relación lineal con el logaritmo de la edad. La relación puede interpretarse por la hipótesis de que se requieren varias mutaciones para originar un cáncer y que la probabilidad de cada mutación es proporcional a la edad. La pendiente de la línea sugiere cinco mutaciones.

Modelo de carcinogénesis

Se han propuesto varios modelos para explicar la acumulación de la frecuencia de cáncer en la edad avanzada. Según uno de los más completos, la célula dispone de varios genes, cada uno de los cuales le impide el desarrollo de un tipo determinado de tumor, y no formará ninguno hasta que cada gen no quede inactivado por mutación. Debido a que las mutaciones pueden introducirse en cualquier momento de la vida de una célula o de sus antecesoras, la probabilidad de que cualquier célula del organismo sufra mutaciones en un gen particular aumenta en relación directa con la edad. La probabilidad de que la célula presente mutaciones en todos los *n* genes protectores, es decir, genes que impiden la formación de tumores (convirtiéndose en cancerosa), se elevará a la *enésima* potencia con la edad. Por lo tanto, el peligro de cáncer aumentará en relación con la edad, hasta su potencia *enésima;* dicho de otra manera, el logaritmo de la incidencia de cáncer deberá guardar una relación lineal con el logaritmo de la edad. La teoría suele ir en perfecto acuerdo con la distribución observada en gráficas estadísticas de la incidencia del cáncer por edades. Suponiendo que el modelo sea correcto, puede calcularse, a partir de la pendiente de la relación logarítmica lineal, el número de mutaciones necesarias para dar origen a un cáncer. En un cáncer típico, como el del intestino grueso, el número parece ser cinco, aproximadamente.

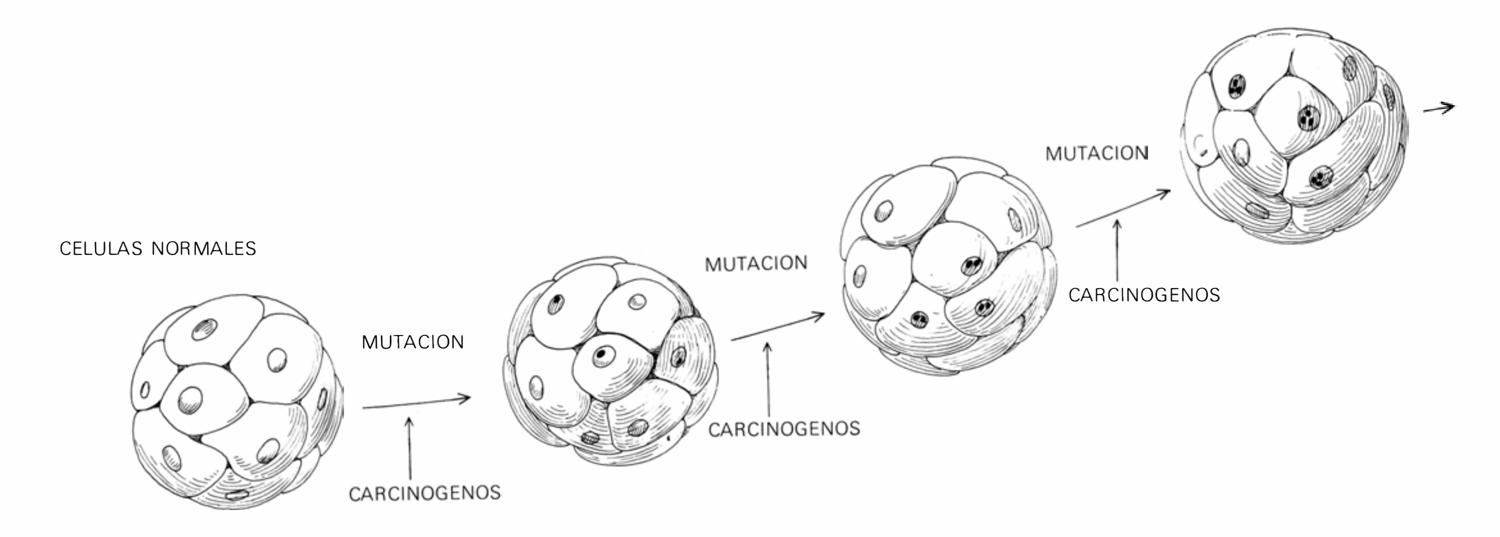
Esa interpretación de la relación entre

edad e incidencia tiene una implicación importante. Se considera a cada tipo de cáncer como resultado final de distintas mutaciones que pueden haber ocurrido en cualquier momento de la vida del paciente. Por tanto, el período de incubación total de cualquier tipo de cáncer se remonta al momento en que tuvo lugar la primera mutación, de modo que cubre a menudo gran parte de la vida del paciente. Ello puede demostrarse en algunos casos. Por ejemplo, la incidencia del cáncer de pulmón en varios países no es directamente proporcional al número de cigarrillos que sus habitantes fuman hoy, sino al número que fumaban hace 20 años; de forma parecida, los tumores profesionales inducidos por la exposición a ciertos productos químicos industriales pueden tardar en aparecer hasta 10 y 20 años después que el individuo hubiera abandonado el trabajo. Otro ejemplo particularmente válido, aunque sea poco frecuente, es el cáncer del pene. Esta enfermedad se manifiesta en hombres de avanzada edad pero viene causado, sin duda, por factores que actuaban durante la juventud, ya que puede prevenirse por la circuncisión en los primeros días de la vida, pero no cuando la circuncisión se pospone unos pocos años. Por último, como tendremos ocasión de ver, varios estudios de poblaciones nómadas muestran que la incidencia de muchos tipos de cánceres comunes se halla determinada parcialmente por el ambiente en que discurre la juventud. De todo lo cual se deduce que cuando indaguemos las causas de un tipo particular de cáncer no debemos polarizar exclusivamente la atención en el pasado reciente del paciente. A la inversa, y ello debe destacarse, si una población se expone, sin conciencia de ello, a un agente cancerígeno, pueden transcurrir muchos años antes de que afloren las primeras manifestaciones de la enfermedad y puede ser entonces demasiado tarde para hacer frente a la ola de tumores que se avecina.

Factores ambientales

El conocimiento de la relación entre la edad y la tasa de muerte no precisa las causas que determinan un cáncer; informa sólo de que los pasos que a él conducen probablemente se han ido acumulando durante muchos años. Deseamos saber si las causas o factores principales son ambientales (y, por tanto, potencialmente evitables). Por ejemplo, si fueran mutaciones las distintas etapas que jalonan la formación de un cáncer, habría que saber hasta qué punto vienen inducidas por mutágenos externos, del ambiente, o bien se originan por error espontáneo en la replicación del ADN. Para ahondar en esa distribución, resulta útil observar el influjo que las migraciones de una comunidad de un país a otro tienen sobre la incidencia de cáncer. Se han estudiado muchas poblaciones; todas ellas muestran que el ambiente representa un papel decisivo. Verbigracia, el cáncer de estómago es mucho más común en Japón que en Estados Unidos, al revés de lo que ocurre con el cáncer de intestino grueso, de mama y de próstata. Cuando un japonés emigra a Norteamérica tales diferencias se pierden en una o dos generaciones *(véase la ilustración superior de la página 97)*. Como quiera que los inmigrantes japoneses, y sus hijos, tienden a casarse dentro del propio grupo, el cambio en la incidencia debe atribuirse a un cambio en el ambiente más que a factores genéticos; más aún, dado que la incidencia de los diferentes tipos de cáncer tarda más de una generación en alcanzar los niveles típicos de Estados Unidos, algunos agentes causales habrán de ser factores del tipo de la dieta, que tiende a persistir como parte de una herencia cultural, mejor que factores del tipo de la polución atmosférica, que tiende a ser idéntica para toda la población de un lugar determinado. De modo similar, los judíos que emigran a Israel desde Europa o Estados Unidos tienen una incidencia de cáncer típica de su país de origen, pero sus hijos, nacidos en Israel, cuentan ya una tasa muy inferior en casi todos los tipos de cáncer. A este respecto, esa primera generación se asemeja a la población judía y árabe indígenas y a los inmigrantes judíos que proceden de Asia y Africa.

Dentro incluso de un mismo país se puede detectar la influencia de variaciones locales del ambiente y de las circunstancias. En Estados Unidos, el índice de mortalidad para la mayoría de los tipos de cáncer comunes es muy inferior en los titulados de grado superior que en los no titulados; se exceptúan el cáncer de mama y el de próstata, superior entre los titulados. Siguiendo en Norteamérica, los individuos de piel muy clara tienen mayor probabilidad de mo-



UNA TEORIA DE LA CARCINOGENESIS establece que varios genes independientes impiden que las células se conviertan en cancerosas; los tumores sólo se desarrollan cuando se acumulan mutaciones en todos esos genes dentro de una misma línea celular. Las mutaciones

rir por cáncer de piel si viven en uno de los soleados estados del Sur; mientras que en la zona de las Montañas Rocosas la muerte por cualquier tipo de tumor es mucho menor que el promedio nacional. Las diferencias en las tasas de mortalidad en el caso norteamericano suelen ser menos del doble, pero este estrecho margen implica que una proporción sustancial del índice de mortalidad podría prevenirse mediante el control de los factores que caben imputar al ambiente.

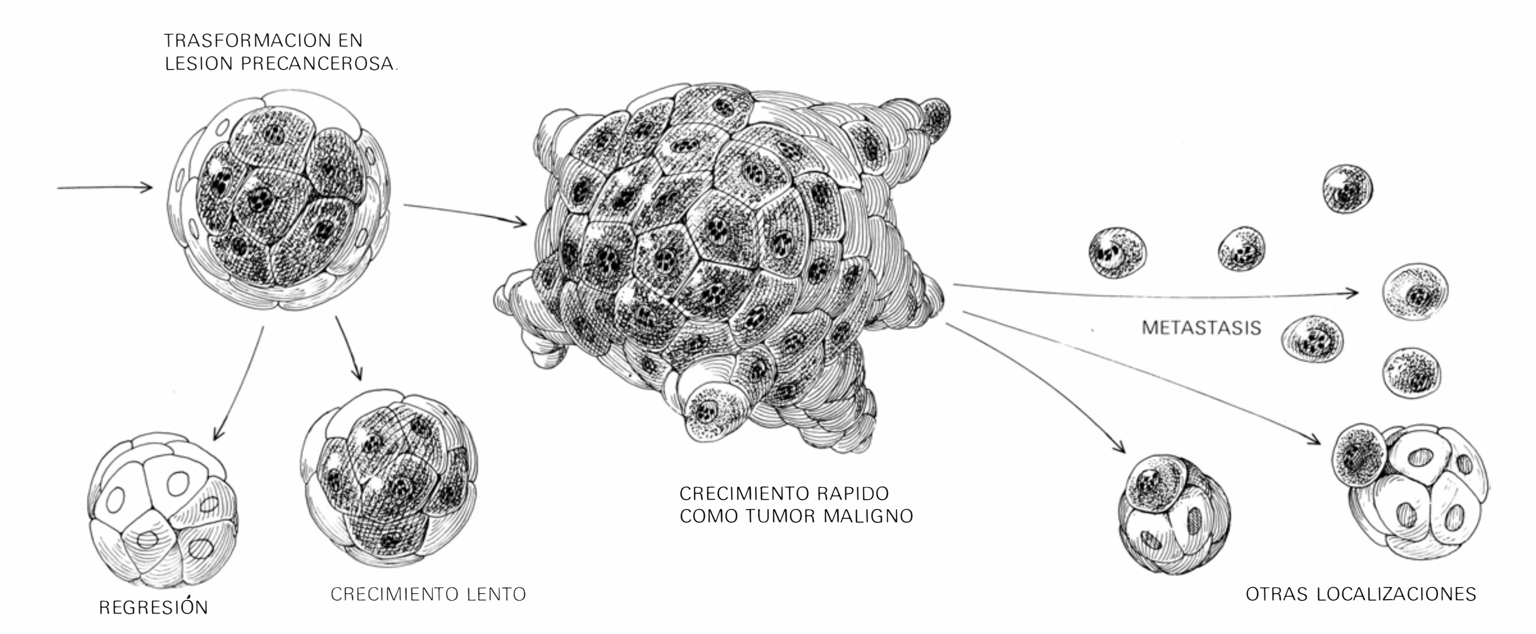
Incidencia y causas

Al demostrar la influencia del ambiente no queda excluida, por supuesto, la posibilidad de una contribución genética a la carcinogénesis. Resulta de enorme interés conocer si los factores genéticos son importantes en la formación de los tipos de cáncer más comunes; y ello aunque sólo fuera porque los programas de detección, para conseguir el diagnóstico precoz, costarían muchísimo menos al poderse acotar de antemano un sector de la población particularmente expuesto. Se sabe que ciertas enfermedades hereditarias, raras, se hallan asociadas a un riesgo altamente elevado de contraer determinados tipos de cáncer; por ejemplo, un defecto hereditario en los enzimas que reparan el ADN dañado por la luz ultravioleta, denominado *Xeroderma pigmentosum,* conduce a múltiples carcinomas de la piel. Algunos tipos de cáncer raros en los niños parecen causados por mutaciones hereditarias. Sin embargo, aun considerándolos conjuntamente, éstos y otros tipos de cáncer, claramente congénitos, son demasiado raros como para contribuir de una manera destacada al problema de la aparición de los tumores.

No es fácil determinar si la frecuencia de los tipos comunes de cáncer se halla influida o no, significativamente, por factores genéticos. Los familiares más allegados comparten el mismo ambiente y, por tanto, podrían mostrar alguna tendencia hacia los mismos tipos de cáncer, aun en ausencia de una determinación genética. Entre las formas más comunes, la que tiene una propensión mayor a agruparse en familias es el cáncer de mama, que duplica, aproximadamente, su frecuencia normal entre las pacientes allegadas a las que ya lo sufren. Tales efectos son menos obvios en los otros tipos. La demostración más palmaria de que la herencia influya en el cáncer sería la constatación de que los gemelos homocigóticos fueran más propensos a sufrir el mismo tipo de cáncer que los gemelos heterocigóticos del mismo sexo. Mas, por tomar un solo ejemplo, un estudio de los gemelos nacidos en Dinamarca a partir de 1870 ni siquiera ha podido establecer este simple hecho. Por tanto, la contribución genética no es, probablemente, una variable importante, al menos cuando se considera el problema del cáncer a escala nacional.

El fin último de los estudios epidemiológicos es la prevención del cáncer mediante la identificación de sus causas y la remoción de las mismas. Ello comporta múltiples y variados problemas, como se desprende de lo que ya conocemos respecto a las causas de cuatro tipos de cáncer: los carcinomas de pulmón, de estómago y de intestino grueso y las diferentes formas de leucemia y linfomas.

El cáncer de pulmón es una enfermedad típica del siglo veinte *(véase la ilustración de la página 93).* Al principio, afectaba únicamente a los varones, pero desde hace poco se localiza también en las mujeres. En Estados Unidos contabiliza casi un tercio de la mortalidad absoluta por cáncer en varones y, en Inglaterra, la mitad aproximadamente. Desde el principio, se creyó que la causa más probable era el consumo de cigarrillos; después de todo, a esa nueva forma de polución atmosférica estaban expuestos los hombres, no las mujeres. Sin embargo, esta explicación no acababa de satisfacer. La incidencia del cáncer de pulmón en diferentes países, podía objetarse, no guarda una relación simple con su consumo de cigarrillos per cápita. La mayoría de las dificultades se disiparon cuando se comprobó que el período de incubación era muy largo. Quedan todavía por despejar muchas incógnitas, pero el hecho fundamental no ofrece ya dudas. El fumador aumenta su riesgo de muerte por carcinoma pulmonar de diez a cincuenta veces, siendo el valor exacto una función de la cantidad fumada, del país en que vive y otros factores distintos. Cuando muchos miembros de un grupo dejan de fumar, la mortalidad por carcinoma pulmonar del grupo, tomado como una unidad, declinará. Hay motivos suficientes para suponer que la abolición del vicio del tabaco erradicaría en



(puntos negros dentro de los núcleos celulares) **raramente son espontáneas, antes bien parecen causadas por factores carcinógenos ambientales.**

Formada ya una lesión precancerosa, en muchos casos remitirá o avanzará lentamente; muy pocas lesiones llegarán a la invasión metastática.

gran parte el cáncer pulmonar, la forma más común de muerte por cáncer. No obstante, no se ve por ahora ningún indicio de que pueda prohibirse el uso del tabaco. Las clases profesionales fuman menos que antes pero las pobres fuman más. Podría argüirse con razón que muy pocas sociedades occidentales podrían permitirse la abolición de un hábito que da vida a una gran industria secundaria, genera substanciosas ganancias y mata, principalmente, a los miembros de edad más avanzada de la población, quienes, de otra manera, resultarían gravosos para la beneficencia gubernamental y la seguridad social.

El cambio en la incidencia del cáncer de estómago ha sido tan espectacular como el ocurrido con el cáncer pulmonar, pero de sentido contrario. En Estados Unidos, la mortalidad por cáncer de estómago ha reducido su índice en ocho veces, aproximadamente, a lo largo de los últimos cincuenta años. De lo cual podemos concluir que un factor único o un grupo de factores estrechamente relacionados tuvieron que ser los responsables de la mayoría de los casos observados en el pasado; y ello por la sencilla razón de que, si hubieran existido muchas causas no relacionadas, difícilmente se hubieran inhibido todas a la vez. Aunque sea por casualidad, el cáncer de estómago ha visto así reducida su frecuencia.

El cáncer de intestino grueso no ha registrado grandes cambios de incidencia en el tiempo, pero varía grandemente de un país a otro, variación que debería proporcionar alguna clave referente a su causa; en general, la incidencia es directamente proporcional a la riqueza del país. Entre las causas a indagar parece obligado detenerse en la dieta, al igual que al explicar el cáncer pulmonar resultaba plausible la búsqueda de un carcinógeno inhalado. De entre los agentes causantes, el factor probable es un alto nivel de carne en la dieta o bien un bajo consumo de cereales *(véase la ilustración de la página 99)*. Se ha avanzado un modelo de mecanismo para el desarrollo del cáncer de intestino grueso, según el cual las bacterias intestinales normales convierten varios componentes de la bilis en substancias carcinógenas. La conversión podría intensificarse con las dietas pobres en residuos que, como se sabe frenan el tránsito del contenido intestinal.

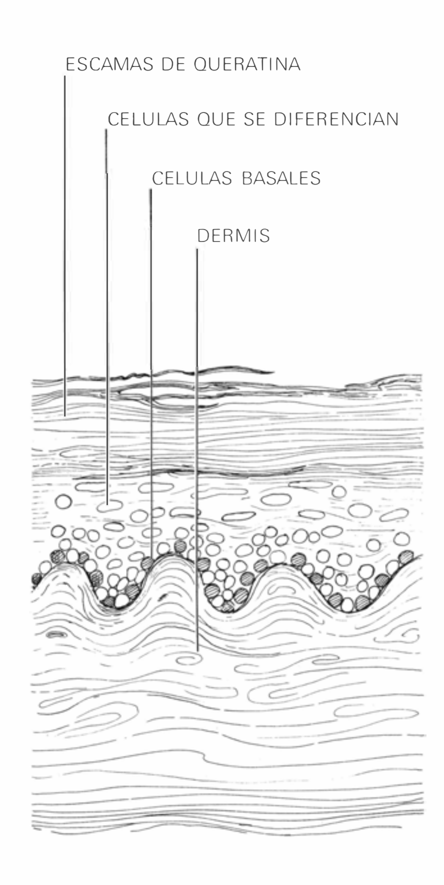
Los carcinomas de pulmón, estómago y de intestino grueso muestran la posibilidad de sacar conclusiones acerca de las causas de un tipo de cáncer, a partir del modo como varía su incidencia de un año a otro o de un país a otro. El cuarto y último ejemplo lo compone un grupo heterogéneo de tipos de cáncer que interesan a las distintas células del sistema inmunitario; comprende las múltiples variedades de leucemia y algunas formas de cáncer de los ganglios linfáticos, entre las cuales la enfermedad de Hodgkin es una de las más frecuentes. Numéricamente son bastante raras, pero adquieren importancia porque suelen afectar a niños y adultos jóvenes. En los últimos tiempos se les ha dedicado mucha atención, porque se cree que pueden estar causadas por virus.

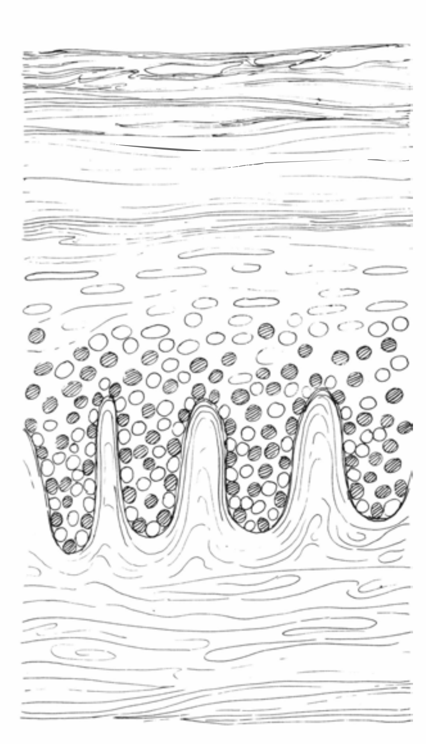
Teoría vírica

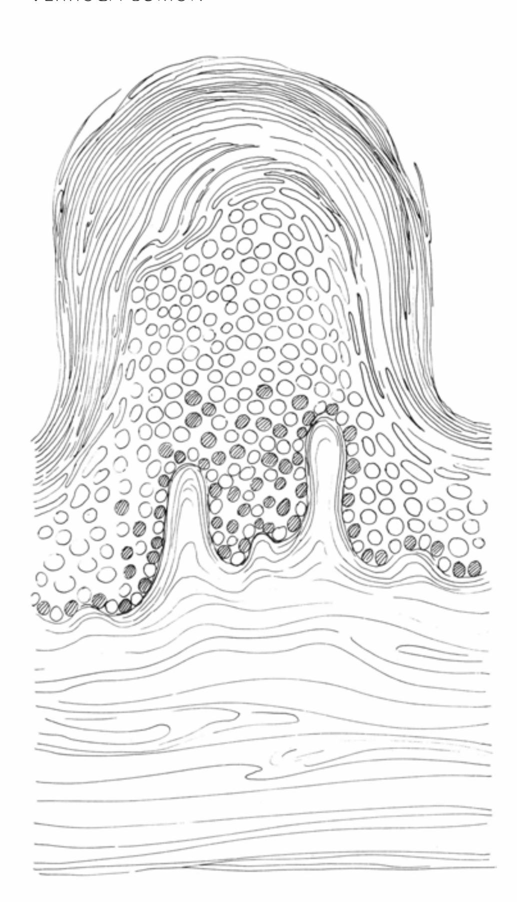
La idea de que los virus puedan hallarse implicados en la carcinogénesis nace de la observación de un hecho experimental: ciertas enfermedades, del tipo de la leucemia de los pollos, gatos y cepas consanguíneas de ratones, pueden inducirse en animales jóvenes inoculándoles virus extraídos de animales leucémicos. La idea de que la leucemia humana, y quizás algunos otros tipos de cáncer humano, puedan producirse también por virus, resulta atrayente por dos razones. Primera: la inducción vírica del cáncer es un fenómeno que puede estudiarse con la metodología de la biología molecular, en tanto que la carcinogénesis debida a mutágenos químicos comporta una investigación más compleja. Segunda: existe la esperanza de que, una vez establecido que un virus sea la causa de un cáncer humano dado, estemos a un paso de su vacuna y, por tanto, de prevenir el tumor en cuestión. Por desgracia, la relación de la mayoría de virus que producen la leucemia animal con sus huéspedes encierra suma complejidad. Gran parte de estos virus se transmiten de un animal a otro "verticalmente", es decir, por herencia; en tanto que los virus patógenos comunes emigran "horizontalmente". Además, supuesta incluso la presencia del virus, el desarrollo de la leucemia en el animal adulto requiere algún fenómeno desencadenante independiente del virus. Por ejemplo, el virus asociado con la leucemia del ratón se hereda por muchas cepas salvajes de ratón y está presente en otras muchas cepas consanguíneas, pero la leucemia se da probablemente muy poco entre las cepas salvajes y puede prevenirse en las cepas consanguíneas merced a una ligera restricción en la dieta; por consiguiente, parece que el cáncer de sangre viene desencadenado, a la vez, por factores dietéticos y por acción vírica. Cuando se adquiere el virus por transmisión horizontal, solamente unos pocos animales desarrollan la leucemia; por ejemplo, el virus leucémico de los gatos se difunde horizontalmente, pero, al parecer, únicamente causa la leucemia en el caso en que el gato sea infectado por una dosis vírica muy alta y no disponga de la respuesta inmunitaria normal. La causa desencadenante puede ser, por tanto, algún otro fenómeno que deprima el sistema inmunitario.

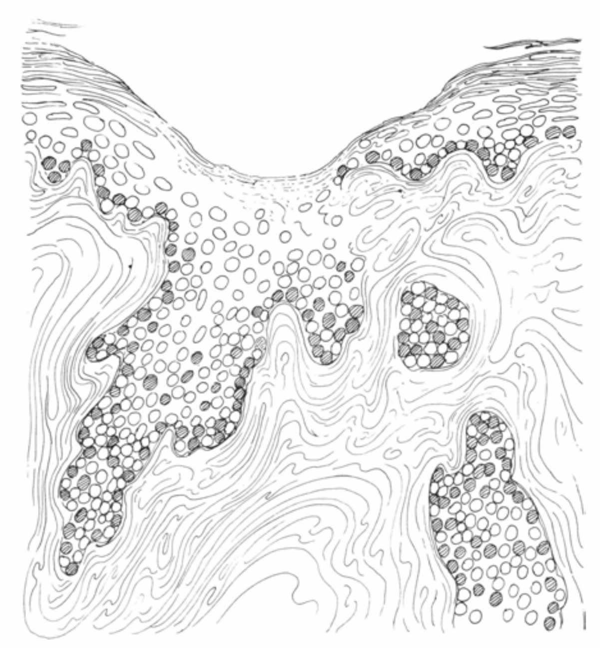
La búsqueda del virus que produce cáncer en los humanos se ha desarrollado a distintos niveles. Unos investigadores estudiaron los virus que suelen hallarse asociados a tumores particulares; otros consideraron la frecuencia familiar de ciertas enfermedades del tipo de la leucemia infantil, que podría indicarnos el carácter hereditario de la leucemia vírica; otros llevaron a cabo investigaciones estadísticas en torno a la incidencia, en determinados lugares y tiempos, de casos de leucemia y enfermedades afines que pudiera orientar sobre la transmisión horizontal de un agente infeccioso. Hasta ahora no hay pruebas concluyentes sobre la existencia de una clase de cáncer humano causado regularmente por acción vírica. Algunos tipos de cáncer se hallan asociados frecuentemente a niveles muy altos de anticuerpos contra ciertos virus; se ha demostrado, asimismo, que algunas células cancerosas contienen ácido nucleico vírico y que determinadas formas que ocurren en los ganglios linfáticos (la enfermedad de Hodgkin y el linfoma de Burkitt, un cáncer raro de los niños del Africa tropical) se presentan, ocasionalmente, de manera casi epidémica. En cada ejemplo, media siempre algún factor distinto que complica y hace difícil la interpretación. Hemos de suponer que tarde o temprano se encontrará algún virus que induzca cáncer específicamente en el hombre, por la sencilla razón de que conocemos la existencia de virus que inducen cáncer en los animales. Conviene tener presente que, con pocas excepciones, los virus producen el cáncer en los animales sólo cuando éstos son muy jóvenes y la inoculación se realiza en combinaciones especiales, mientras que cualquier animal, de cualquier edad, sufrirá cáncer cuando se le administre el carcinógeno químico adecuado por cualquier vía.

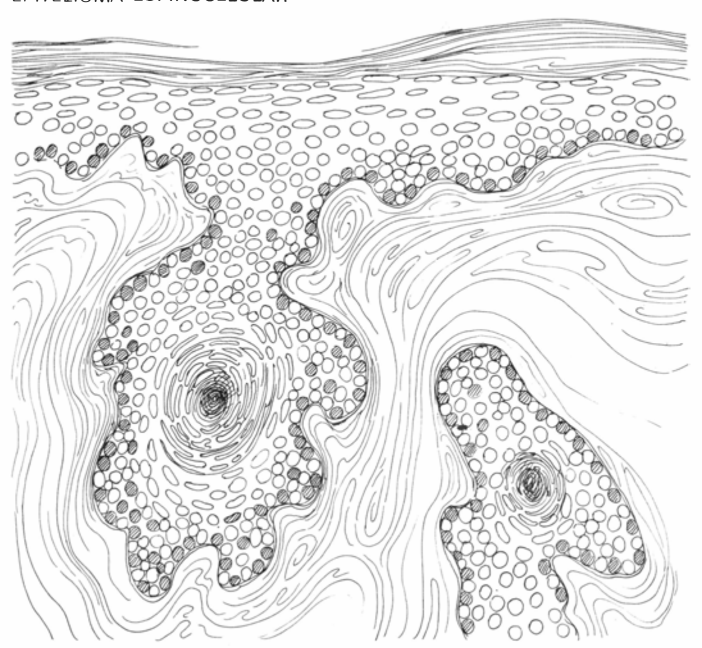
Esta breve revisión de la epidemiología del cáncer humano pretende mostrar lo que sabemos y lo que esperamos encontrar sobre las causas del cáncer. A fin de poner de relieve que el objetivo que se persigue es de orden práctico, a saber, la prevención de la pérdida de vida eficaz por culpa del cáncer, se han elegido









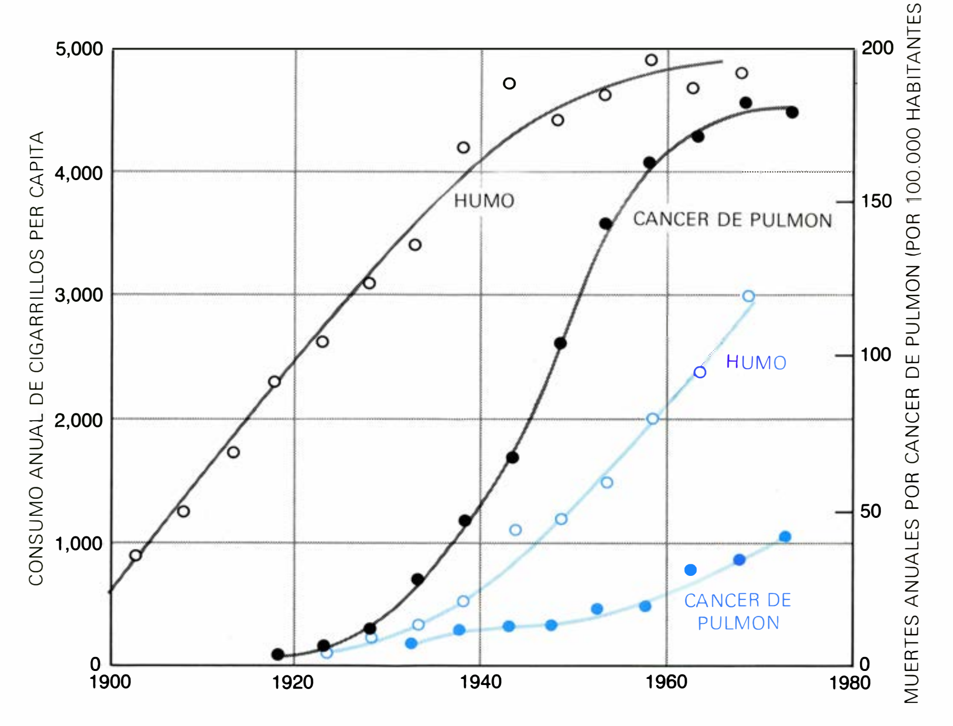


LAS ENFERMEDADES DE LA PIEL que comportan anomalías del control del crecimiento engloban tumores malignos y otras enfermedades consideradas menos complicadas. En la piel normal, la capa más exterior, la epidermis, está formada por un solo estrato de células basales recubierto por otras células en vías a la diferenciación para formar escamas o láminas planas de queratina. Sólo las células de la capa basal son capaces de división *(sombreado)*. En las psoriasis, se halla incrementado el número de células basales en división, de modo que se producen células a una velocidad superior a la que se desprenden de la superficie; y así, la epidermis se engruesa. En la verruga común, el proceso de diferenciación sufre una deceleración, por lo que las células se acumularán en cada estadio de su ciclo vital lo que determina también un espesamiento. Ninguna condición de las dos es cancerosa y, en ambas, la organización global de la piel es normal; en particular, permanece bien definida la frontera que separa la dermis de la epidermis. En los carcinomas cutáneos, la organización de los tejidos está perturbada y las células trascienden los límites territoriales normales. El epitelioma basocelular viene provocado por células basales que siguen multiplicándose –no diferenciándose– e invadiendo la dermis. El carcinoma de células escamosas está compuesto por células que se diferencian con relativa normalidad, pero fuera de la epidermis. Las cuatro enfermedades parecen estar causadas por fallos en la comunicación intercelular.

los ejemplos entre los distintos tipos de cáncer más comunes en la sociedad occidental, y esa elección ha pretendido destacar la importancia de la dieta y de ciertos hábitos como el de fumar. Cada grupo de población del mundo padece sus tipos de cáncer peculiares. Si se tomaran las medidas oportunas de salud pública, de forma que pudiera reducirse la incidencia de cada tipo de cáncer a sus niveles mínimos observados en todo el mundo, la incidencia total de las enfermedades cancerosas se vería reducida a menos de una décima parte. Lo cual vendría a equivaler a la reducción en la mortalidad por enfermedades infecciosas que se ha conseguido en los últimos cincuenta años.

Mecanismo de carcinogénesis

La epidemiología se ocupa de las causas y, por tanto, de las perspectivas para su prevención. El problema del cáncer puede afrontarse también desde la óptica de los mecanismos y las perspectivas de curación. Hasta hoy no disponemos de una terapia general eficaz y no hay indicios de que pueda llegarse a ella pronto. Unos cuantos tipos de cáncer (entre ellos, la enfermedad de Hodgkin y el linfoma de Burkitt) pueden curarse frecuentemente por una combinación de drogas citotóxicas; asimismo el crecimiento de formas que atacan a órganos sensibles a las hormonas sexuales puede retardarse con la administración de tales hormonas. Mas para la gran mayoría de tipos no existe una droga específica y deben tratarse por una combinación adecuada de cirugía, radiaciones y drogas citotóxicas, procedimiento que ha demostrado ser, empíricamente, el más indicado a tenor de los resultados. Pero éstos no son nada optimistas. No llegan ni a la mitad los pacientes que logran sobrevivir cinco años a partir del momento en que se verificó en ellos el primer diagnóstico. La muerte sobreviene casi siempre por metástasis, es decir, por diseminación de las células cancerosas a otros sitios lejanos. Si no fuese por tal diseminación, muy pocas formas quedarían fuera del alcance de la cirugía moderna. Se observa, en efecto, que muchos tumores benignos crecen rápidamente y alcanzan grandes tamaños, a pesar de lo cual son raramente letales. En cuanto causa de mortandad, el cáncer se halla más cerca de un defecto del mecanismo que establece normalmente los límites territoriales de las células que de una anormalidad del control de crecimiento propiamente dicho. Al ser, predominantemente, una enfermedad de los epitelios, conviene examinar su biología a nivel de los sistemas de control que han de determinar la forma y establecer los límites territoriales de las células epiteliales. A este propósito, la piel es un órgano que sirve como modelo para la comparación del comportamiento de las células normales con el de las células cancerosas.



EL CONSUMO DE TABACO Y EL CANCER DE PULMON están indiscutiblemente relacionados, pero la naturaleza de esa relación permanece oscura debido al largo período de latencia que transcurre entre el aumento en consumo de cigarrillos y el aumento en la incidencia del cáncer de pulmón. Los datos corresponden a Inglaterra y Gales. Entre los hombres *(en negro)* empezó a aumentar el hábito de fumar a principios de siglo, pero la correspondiente tendencia en la mortalidad por cáncer no comenzó hasta después de 1920. Las mujeres *(en color)*, se habituaron más tarde; en proporción inferior, se registra en ellas ya la incidencia de cáncer.

El epitelio de la piel se llama epidermis; si hacemos un corte, observaremos su carácter estratificado, con un espesor aproximado de cinco a diez células, que se apoya en otra capa celular menos densa con función de sostén, la dermis. La epidermis se regenera continuamente a sí misma gracias a la división de las células que ocurre en su capa más profunda o basal, la cual está en contacto con la dermis. En virtud de esa constante división celular, hay un continuo desplazamiento de células hacia las capas más superficiales. Allí comienzan a diferenciarse según un programa preestablecido: se achatan, inician la síntesis de una proteína insoluble denominada queratina y pierden sus núcleos. Finalmente, se funden en láminas, las escamas de la piel, que se irán desprendiendo de la superficie. Como resultado de este programa de desarrollo, el organismo se halla separado del entorno ambiental por una capa relativamente impenetrable de queratina insoluble que se desprende y se substituye sin solución de continuidad.

Del comportamiento de las células epidérmicas podemos deducir que tienen que estar sometidas a distintos tipos de control. Primero, el hecho de que las únicas células que se dividan sean las que se encuentran en contacto con la dermis subyacente sugiere que hay algún tipo de señales de corto alcance que comunica a la dermis con las células basales; en ausencia de tales señales, la célula epidérmica deja de multiplicarse y empieza a diferenciarse. Segundo, para impedir que las células basales que se multiplican invadan la dermis, algún mecanismo debe establecer y controlar la frontera que limita ambas capas. Tercero, tiene que haber algún sistema de señales laterales que regule el espacio que separa las estructuras epidérmicas, cuales son los folículos pilosos y las glándulas sudoríparas.

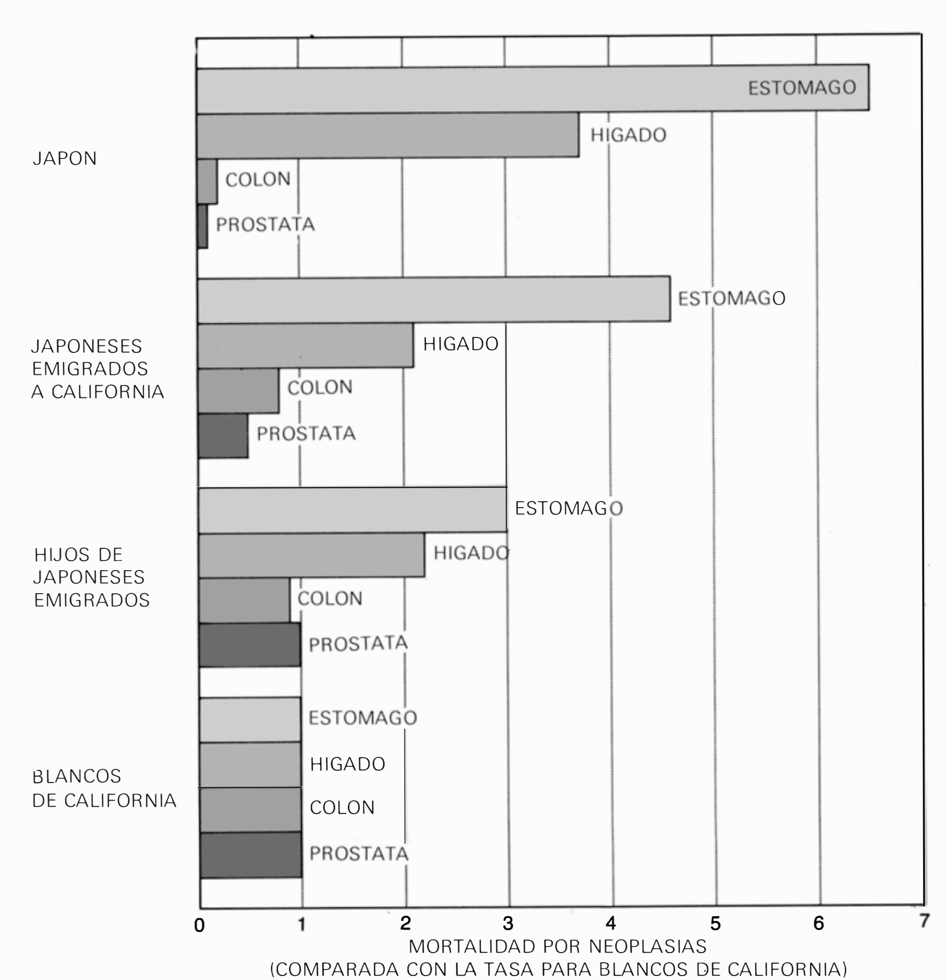
Además de la regulación local del crecimiento epidérmico, el comportamiento de las células reclama la existencia de otros sistemas de control general. Aunque algunas características superficiales, como las huellas dactilares, se expresan en la epidermis, están determinadas por la dermis; si se extirpa tejido epidérmico del muslo, por ejemplo, y se injerta en la palma, adquirirá un mayor grosor y seguirá las líneas características de la palma en cuestión. Si una zona de la piel se halla sometida a un creciente

desgaste, se modifica el programa de diferenciación, de forma que encuentra la profundidad de las capas celulares y se endurece la capa de queratina, formando un callo. Cuando se extirpa de una zona la epidermis, vuelve a ser colonizada gracias a un aumento de la tasa de división celular de la epidermis vecina. Al implantar una porción de piel en un tejido subcutáneo, la epidermis degenera.

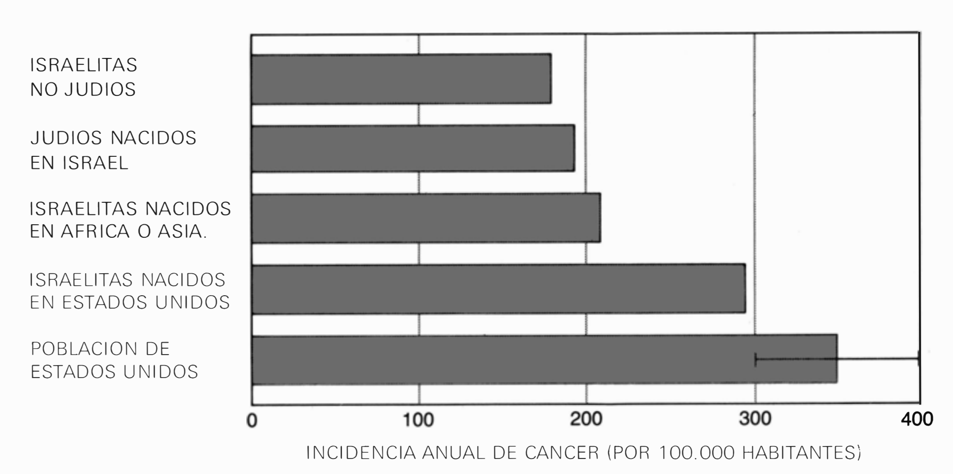
No se conocen los mecanismos de acción de todos esos controles. Implican, necesariamente, una comunicación intercelular; ésta podría llevarse a cabo mediante un gradiente de concentración de una substancia, que se difunda libremente, segregada por una célula y detectada por otra; o bien podría exigirse el contacto intercelular directo. Sea cual fuere el mecanismo, la comunicación podría interrumpirse por un defecto en la célula emisora o en la célula receptora, y puesto que, al parecer, existen muchos sistema de señales en la piel, cada uno de los cuales opera de una forma más o menos autónoma, tendrían que darse muchas alteraciones distintas en el control del crecimiento. Que es exactamente lo que ocurre.

Dos enfermedades cutáneas, perfectamente conocidas, que representan anormalidades no cancerosas en el control del crecimiento son la psoriasis y la verruga común. En la psoriasis aumenta el número de células basales que se multiplican, de modo que la capa celular basal llega a tener el espesor de unas 10 células. Por lo que se sabe, esta enfermedad se debe a un fallo en la comunicación entre la dermis y las células basales. La verruga común, causada por una infección vírica, es un engrosamiento local de todas las capas de la epidermis. En esta segunda enfermedad, la diferenciación de las células epidérmicas parece sufrir una remisión drástica, de suerte que en cada nivel de diferenciación se encuentra mayor número de células; pero la disposición global de las células conserva su ordenación precisa y la frontera con la dermis permanece inalterada.

Hay dos tipos de cáncer comunes de la epidermis: el carcinoma basocelular y el carcinoma espinocelular. El carcinoma basocelular está compuesto por células que proceden de la capa basal, las cuales parecen haber escapado al control del sistema que se ocupa de mantener la frontera entre la dermis y la epidermis. Las células cancerosas invaden la dermis y los tejidos subyacentes, formando una úlcera irregular y erosiva; a esta forma de cáncer cutáneo se le da, a veces, el



LOS CAMBIOS DE INCIDIENCIA de distintos tumores malignos, en los emigrantes de Japón a Estados Unidos, prueban que ellos se hallan causados por componentes del ambiente que difieren en uno y otro país. La incidencia respectiva se expresa por el cociente de la mortalidad en la población considerada con respecto a una población hipotética de blancos de California con la misma distribución por edades; se asigna valor uno a la mortalidad blanca. Las tasas de muerte entre emigrantes e hijos de emigrantes tienden irreversiblemente hacia la norma de California, pero el cambio requiere más de una generación, probablemente porque los agentes causales sean factores ligados a hábitos de vida y no fenómenos del azar a los cuales todos están necesariamente expuestos por igual.



LOS FACTORES AMBIENTALES explican también las diferencias entre Israel y Estados Unidos respecto a la incidencia global del cáncer. La frecuencia del cáncer en Israel es relativamente baja, y es la misma en la población judía y no judía (ésta principalmente árabe); la incidencia en Estados Unidos, aunque varíe significativamente de una a otra región, resulta casi el doble. La alta incidencia entre los israelitas oriundos de Occidente depende quizá del largo período de latencia de los tipos de cáncer: las personas expuestas a carcinógenos en la juventud pueden desarrollar tumores más tarde, incluso si se han desplazado a un ambiente cuyos niveles de carcinógenos sean más bajos.

nombre de ulcus rodeus. A pesar de su gran poder de invasión local, el carcinoma basocelular no suele metastatizar casi nunca, lo cual permite suponer que las células cancerosas requieren también alguna señal de la dermis para multiplicarse.

El segundo cáncer de piel, el carcinoma espinocelular, consiste igualmente en grupos desordenados de células; mas, a diferencia de las del cáncer basocelular, aquéllas sufren una diferenciación casi normal en escamas de queratina. El carcinoma espinocelular es espacialmente menos invasor, pero puede dar origen a metástasis distantes. Sus células, pues, poseen verosímilmente propiedades de cohesión propias del tejido epidérmico normal, si bien su relación de dependencia de las señales de las dermis es menor.

Se registran otros muchos desórdenes en el crecimiento de la epidermis, que reflejan las distintas maneras como puede perturbarse la función de los sistemas reguladores. Resulta significativo que sólo cuando las células se liberan de las imposiciones de la territorialidad pueden formar un cáncer potencialmente letal. Además, como ilustran los dos tipos de cáncer comunes de la piel, la territorialidad puede perderse de varias maneras, y cada clase de pérdida tendrá un efecto distinto sobre el comportamiento de las células. Por lo tanto, el cáncer no puede considerarse una enfermedad única, determinada en cada caso por una misma disfunción celular. A pesar de que los carcinomas basocelulares y espinocelulares se originan ambos en la epidermis y hallan su causa común en la luz solar, las células afectadas se desenvuelven distintamente.

La respuesta inmunitaria

No podremos esperar comprender el comportamiento de un cáncer mientras no seamos capaces de discernir las señales que pasan entre las células y las fuerzas que separan grupos de células diferentes en los organismos multicelulares. Aunque se sabe poco de los sistemas de comunicación intercelular, es evidente que las moléculas de las superficies celulares deben determinar las señales que la célula puede recibir y cómo interacciona la célula con otras inmediatas. Algunas anomalías de las células cancerosas tienen que expresarse por tanto en su superficie.

En la superficie de cada célula humana se encuentran moléculas denominadas antígenos de histocompatibilidad, que la distinguen de las células de los otros individuos. Forma parte de las misiones de nuestro sistema inmunitario destruir las células que lleven antígenos de histocompatibilidad equivocados. Es natural preguntarse si puede encontrarse algún medio para movilizar al sistema inmunitario contra cualquier componente extraño que hubiera en la superficie de las células cancerosas. Desde el descubrimiento de la inmunidad celular, a comienzos de siglo, se ha repetido una y otra vez la hipótesis de que una de las funciones principales del sistema inmunitario era destruir los tumores y, por consiguiente, que la aparición del cáncer en la edad avanzada muestra sencillamente el debilitamiento progresivo del sistema inmunitario.

Al igual que ocurre con la teoría vírica del cáncer, la idea es atractiva porque ofrece la esperanza de una curación, en este caso por alguna forma de inmunoterapia, sin tener que aguardar ulteriores avances en la biología fundamental. Por desgracia, los hechos desmienten la teoría.

Al tratar a los pacientes con inmunodepresores para que toleren el trasplante de un órgano, y al examinar individuos con defectos hereditarios en el sistema inmunitario, se ha observado un notable incremento en la incidencia de ciertos tipos raros de cáncer de las células del sistema inmunitario, pero no en la mortalidad por cualquiera de los tipos comunes. En casos de avanzada ancianidad, si se comprueba la reactividad de estas personas ante antígenos adecuados, la mortalidad anual se encuentra más elevada en los que presenten la reactividad deprimida, pero la muerte no está causada por cáncer. De modo parecido, ratones con diversos trastornos en el sistema inmunitario, inducidos o heredados, pueden mostrar un incremento en la susceptibilidad a ciertos virus cancerígenos, pero no en la susceptibilidad a los carcinógenos químicos, ni tampoco aumenta la frecuencia de cáncer espontáneo.

Por otra parte, no hay duda alguna de que las células de muchos tipos de cáncer en animales de laboratorio, y de determinados tipos de cáncer en humanos, presentan realmente componentes superficiales anormales que sugieren una respuesta limitada por parte del sistema inmunitario. Se ha planteado así la hipótesis alternativa de que una reacción inmunitaria moderada, a cargo del huésped, estimula la multiplicación de las células cancerosas; en otras palabras, puede existir durante el crecimiento del cáncer cierta selección de variantes celulares dotadas de componentes superficiales nuevos, que provocan una respuesta inmunitaria. Lo cual se opone diametralmente a la idea original de que el sistema inmunitario inhibe el desarrollo de los distintos tipos de cáncer. A pesar de ello, no se descarta en absoluto que algún método de inmunoterapia pueda llegar a dirigirse contra cualquier tipo de anormalidad en la superficie de la célula cancerosa que sea responsable de su comportamiento desorganizado e invasor.

Diagnóstico precoz

Un último aspecto del problema del cáncer que debe estudiarse vuelve a hacer referencia a la evolución de la enfermedad. Desde el punto de vista del paciente, la cuestión más importante es que el cáncer se detecte y extirpe antes de que haya engendrado la metástasis. Se están haciendo grandes esfuerzos para desarrollar programas de diagnóstico muy precoz para los tipos de cáncer comunes, en particular el cáncer de mama y de útero. Por desgracia, cabe presumir un éxito limitado para estos programas.

Para que un cáncer forme tumores secundarios en un órgano distante, no basta que sus células penetren en el torrente sanguíneo o en el sistema linfático por invasión directa desde el tumor primario; se precisa también que las células sobrevivan y se multipliquen en un ambiente extraño. Hay razones para creer que ese último obstáculo sea, con frecuencia, el factor que limita la expansión de un tumor. Las metástasis de los tumores no se implantan al azar. El cáncer de pulmón tiende a colonizar el cerebro y las glándulas suprarrenales; el cáncer de la próstata suele invadir los huesos; una variedad del cáncer de estómago se extiende casi exclusivamente por los ovarios. En otras palabras, la capacidad de crecer en territorios extraños no es una propiedad universal de los tumores. Algunos, como el carcinoma de células basales de la piel y la mayoría de los tumores cerebrales, casi nunca metastatizan; otras formas pueden producir múltiples metástasis en una fase en que el cáncer primario todavía es demasiado pequeño para poderlo diagnosticar. Es un craso error pensar que la historia natural de todo cáncer consiste en crecer hasta un cierto tamaño, al que sucederán inevitables metástasis. A la hora de planificar programas de detección para distintos tipos de cáncer, habrá que prever que unos no se diseminarán, aunque se dejen sin tratamiento durante largo tiempo, y otros serán letales, por muy precoz

que sea el diagnóstico. En consecuencia, sólo para un grupo limitado, intermedio, la detección aportará algún beneficio. Por mencionar un ejemplo muy claro: un programa de control cuidadosamente realizado en Nueva York ha reducido aparentemente la mortalidad por cáncer de mama en el grupo examinado en casi un 30 por ciento. No obstante, parece que la mortalidad se ha reducido sólo para mujeres mayores de 50 años, lo cual sugiere que el cáncer de mama en mujeres más jóvenes, que es capaz de metastatizar, lo hace demasiado pronto para que pueda interceptarse. A consecuencia de la campaña de detección anual realizada sobre una muestra de 31.000 mujeres durante cinco años, se estima en 23 el número de vidas salvadas, número que no es probablemente mayor que la mortalidad femenina absoluta por cáncer de pulmón durante el mismo período en el mismo grupo; y éstas podían haberse salvado de haber cesado todas en el uso del tabaco. Otros dos programas que han tenido cierto éxito corresponden al control del cáncer de útero, y, en Japón, al control de cáncer de estómago. En cambio, otro intento de reducir la alta mortalidad por cáncer de pulmón mediante frecuentes exploraciones radiológicas del tórax resultó un fracaso total.

Aparte de su limitada eficacia, los programas de control tienen una complicación adicional: por cada cáncer diagnosticado, revelan otras diez anormalidades aproximadamente, muchas de las cuales parecen ser precancerosas. Mientras no se pueda prever cuál de ellas evolucionará en cáncer, hay que tratarlas todas, lo que supone un problema económico. Está por encima de los recursos de Estados Unidos, o de cualquier otro país, el someter a tratamiento todas las lesiones precancerosas que pudieran detectarse, aun contando con los limitados procedimientos de detección actualmente disponibles, si tales procedimientos se aplicaran a toda la población. La campaña de control que diagnosticó cáncer de piel en el 4 % de la población adulta, rural, del estado de Tennessee encontró lesiones cutáneas precancerosas en la sexta parte de la población; si todas esas lesiones tuvieran que extirparse quirúrgicamente, supondrían un gasto para el estado de varios millones de dólares, y ello teniendo en cuenta que el cáncer de piel es, quizás, la forma que puede diagnosticarse y tratarse con menores costos.

Conclusiones

El problema del cáncer es de interés inmediato y urgente. Todo el mundo espera oír la noticia del descubrimiento de un remedio eficaz y que se le digan las medidas que debe adoptar entre tanto para reducir la probabilidad de su ataque mortal. He intentado hacer un breve análisis de lo que se conoce sobre la naturaleza del cáncer en el hombre y, en particular, sobre las perspectivas para su prevención o tratamiento. Debido a que el cáncer despierta un gran interés y atrae mucho apoyo financiero, los investigadores se ven sometidos a la obligación de poder anunciar alguna novedad terapéutica trascendente o, a falta de ello, algún progreso importante en nuestra comprensión de la enfermedad. En consecuencia, nada más impopular que dar a entender que la deseada curación puede que haya de esperar a que se conozca mucho más sobre las interacciones celulares en organismos muy sencillos. Aunque este punto de vista sea correcto, deberíamos esforzarnos entre tanto por avanzar más en medicina preventiva. Puesto que parece que los programas de detección son de uso limitado y harto costosos, sólo queda la vía de la prevención a través de la búsqueda y erradicación de las causas del cáncer. Esto no ha de considerarse como una invitación a la resignación. Después de todo, a la medicina preventiva se debe, en gran parte, la erradicación de las enfermedades infecciosas.

INCIDENCIA ANUAL (POR 100.000 HABITANTES)
0
10
20
30
40
50
NUEVA ZELANDA
ESTADOS UNIDOS
CANADA
DINAMARCA
GRAN BRETAÑA
SUECIA
HOLANDA
NORUEGA
ISRAEL
ALEMANIA (REP. FED.)
ISLANDIA
ALEMANIA (REP. DEM.)
JAMAICA
FINLANDIA
YUGOSLAVIA
PUERTO RICO
POLONIA
RUMANIA
HUNGRIA
JAPON
CHILE
COLOMBIA
NIGERIA
0
80
160
240
320
CONSUMO DE CARNE (GRAMOS POR PERSONA POR DIA)

LA GEOGRAFIA DEL CANCER sugiere una causa probable de la enfermedad. La frecuencia del cáncer de intestino grueso en la población femenina de 23 países se halla en estrecha relación con el consumo de carne per cápita. Los datos están ajustados para eliminar las diferencias debidas a la edad. Una explicación alternativa atribuye el mismo tumor a un bajo consumo de cereales. Las dos hipótesis –un elevado consumo de carne y un bajo consumo de cereales–, tienden a estar asociados.



Vinos, vides y clima

Tipos de vid muy seleccionados, cultivados adecuadamente en regiones con características climáticas muy precisas, son los factores determinantes de la calidad de los grandes vinos

Philip Wagner

Continuamente está creciendo el número de personas interesadas en saber distinguir entre un buen vino y otro de menor calidad o entre uno óptimo y otro bueno a secas. Esto no sólo sucede en los países del Nuevo Mundo que sólo muy recientemente han empezado a consumirlo de modo habitual, sino también en los países europeos, en los que el vino es elemento tradicional de las comidas de todas las clases sociales desde hace muchos siglos. Frente a esta gran curiosidad, la mayor parte de la gente dispone de pocos conocimientos y un aura de misterio rodea al procedimiento por el que el buen catador califica como corrientes a algunos de los vinos que bebe y, de pronto, manifiesta su atención y respeto hacia otros.

Las opiniones sobre un vino de origen son siempre un tanto subjetivas, en el sentido de que casi todo el mundo tiene sus favoritos. Tales preferencias, sin embargo, se orientan sobre una base definida y objetiva que permite establecer la diferenciación. En lo que sigue intentaremos fijar los elementos fundamentales que permitan tales diferenciaciones. Quien conozca tales principios y posea un sistema sensorial sano y una buena memoria, observará que el vino puede tener un significado mucho mayor de lo que sospechaba. Las puertas están abiertas. La apreciación del vino no es muy distinta a la de la música, la pintura o la literatura.

Fijémonos en un bebedor experto al que no le gusten demasiado los vinos con un regusto dulce. Aquí estriba la subjetividad de su juicio. A pesar de su preferencia (o prejuicio), no dejará de reconocer que un gran sauternes francés o un trockenbeeren alemán es algo muy particular: el resultado de la singular conjunción de ciertas uvas, maduradas en determinadas condiciones, con métodos de elaboración especiales y exigentes. Reconocerá asimismo las diferencias entre tal vino y otro preparado ex profeso para imitarle. En esto consisten las distinciones objetivas.

Quede claro desde un principio que, al hablar de vinos, nos referimos solamente al jugo fermentado de la uva, y no al de fresas, manzanas, ruibarbo o cualquier otra planta. Los diversos tipos de vid llevan todos el nombre del género, *Vitis,* voz latina que significa vid. El género comprende algunas docenas de especies que crecen en diferentes lugares del hemisferio boreal. En su mayoría no tienen interés para la elaboración del vino, aunque algunas tienen otras cualidades importantes para la viticultura contemporánea. El vino que cumple los patrones europeos está elaborado a partir de una sola especie, la vid eurasiática, *Vitis vinifera,* y por algunos híbridos procedentes del cruzamiento de *V. vinifera* con otras especies que producen frutos del tipo de los de *vinifera.*

V. vinifera se subdivide en numerosos grupos o razas, cuya diferencia mutua suele ser muy clara, hasta el punto de resucitar una y otra vez la debatida cuestión de si se trata realmente de una especie pura. Si definimos la especie como algo absoluto, y no como un convencionalismo práctico, habrá que admitir que, a lo largo de la evolución de la vid, ha ocurrido una considerable mezcla de estirpes. Sin embargo, aquí nos basta con afirmar que los grupos y los individuos comprendidos dentro de la expresión *V. vinifera* pertenecen todos a la misma especie general, y que sus diferencias se deben a la selección natural contra las más débiles y a la selección dirigida por el hombre en favor de los dotados de mejores caracteres. Ambos tipos de selección explican la existencia de vinos tan radicalmente distintos entre sí como el retsina y el champán o el oporto y el beaujolais.

Todas las variedades de *V. vinifera* tienen ciertas características comunes. Son caducifolias. Carecen de resistencias a ciertos parásitos, en especial hemípteros, fitófagos del tipo de la filoxera y gusanos del grupo de los nemátodos. Son extremadamente sensibles ante diversas enfermedades fúngicas. Necesitan una buena cantidad de calor y de luz solar para madurar de forma adecuada. Su resistencia al frío invernal es limitada, y su supervivencia se torna incierta donde la isoterma de enero baja muy por debajo de 1º Celsius. Las partes del mundo en donde las vides pueden desarrollarse y producir uva son relativamente restringidas, por más que el área que cubran sea muy grande. La viticultura de las *vinifera* sería tan inviable en las cuencas del Congo y del Amazonas o en las estepas de Saskatchewan y Siberia, como cultivar piña americana en Maryland. El clima tiene la última palabra.

Para dar idea del papel predominante del clima podríamos trazar sobre un mapa de Europa una línea ondulada de oeste a este *(véase ilustración en las dos páginas siguientes).* Europa continúa siendo el continente en donde se cultiva

VIÑEDO DISPUESTO EN BANCALES, plantado recientemente en Napa, California *(fotografía aérea de la página opuesta).* **En la región predomina un clima que corresponde a una variante especial del clima mediterráneo, biestacional, con los inviernos frescos y húmedos y los veranos secos. El valle de Napa recibe la influencia benéfica de la cercana bahía de San Francisco y de la llegada de los vientos fríos del Pacífico, que penetran a través de la Golden Gate. Su altitud le beneficia. Tales factores han determinado que sea la primera región vinícola de California.**

mayor cantidad de uva. La línea empieza en Bretaña, al norte del río Loira, avanza en dirección este y gira luego bruscamente al noreste, cerca de Chartres y encerrando en su parábola a París (existe una pequeña viña en la cumbre de Montmartre). Prosiguiendo hacia el este, la línea pasa por Reims y el norte de la región de Champagne, dobla ligeramente en dirección sudeste hacia Nancy, en Lorena; enfila después en línea recta hacia el norte siguiendo el curso del río Mosela a lo largo de la frontera de Luxemburgo y, de allí, por el tramo mejor conocido del Mosela alemán, hasta Coblenza. Desde allí la línea prosigue con una leve inclinación al sudeste (en realidad dibuja trayectorias a lo largo de un cierto trecho), penetra en Checoslovaquia y, después, junto al río Oder, tuerce bruscamente en dirección sur, llegando a las proximidades de Viena (en la región de los Alpes, donde las vides no pueden crecer, la trayectoria descrita es de trazo interrumpido). Poco después de Viena, se produce otro giro brusco, esta vez hacia el este. Avanza por el norte de la llanura húngara, remonta los Cárpatos, cruza el Dniéper y el Don a cierta distancia del Mar Negro, atraviesa la República Soviética de Georgia por el norte de la cadena del Caúcaso, salta el mar Caspio y acaba perdiéndose en las del Kazakistán.

Esta línea corresponde al límite septentrional del cultivo de la vid eurasiática. Su posición está regulada por dos factores limitantes. El extremo occidental se encuentra bajo la influencia marítima, que atempera el rigor de los inviernos, pero en esta elevada latitud los veranos son excesivamente frescos, por lo que al norte de esta línea los frutos de la vida no alcanzan una maduración suficiente. Más al este, las influencias continentales toman el relevo y, al llegar la línea al Oder, el tiempo invernal resulta ya incompatible con la vid..

No debe considerarse esta línea como algo absoluto, en el sentido de que al norte de ella no puedan crecer cepas. Mediante cruzamientos con otras especies, se consiguen vides de maduración rápida y buena capacidad de resistencia, que perviven en condiciones difíciles y producen vinos aceptables, por lo menos. Estas vides se cultivan en lugares especialmente favorables del sur de Inglaterra e incluso en Moscú, en cuyo recinto de la feria agrícola permanente existe una pequeña viña. Sin embargo, las vides de Moscú son híbridos ultrarresistentes, conseguidos por el naturalista Michurin, sin valor para la viticultura, y aún así hay que arroparlas durante el invierno para que no se mueran. Lo mismo valdría para las partes más frías de la región templada de América del Norte. En suma, la vida eurasiática puede culti-

CLIMA MEDITERRANEO
ESPAÑA
1. SHERRY
2. MALAGA
3. ALICANTE
4. TARRAGONA
5. VALDEPEÑAS
6. ZARAGOZA
7. BAJA RIOJA
FRANCIA
8. ROUSSILON
9. CORBIÈRES
10. MINERVOIS
11. LANGUEDOC
12. CÔTES DE PROVENCE
13. CÔTES DU RHÔNE
14. GAILLAC
ITALIA
15. PIEMONTE
16. TOSCANA
17. REGIONE ALPINA
18. ORVIETO
19. VERDICCHIO
20. VENETO
21. CAMPANIA
YUGOSLAVIA
22. DALMACIA
23. SERVIA
URSS
24. CRIMEA
25. GEORGIA
CLIMA TEMPLADO
ESPAÑA
26. ALTA RIOJA
FRANCIA
27. JURANÇON
28. ARMAGNAC
29. BORDELAIS
a. MEDOC
b. GRAVES
c. SAUTERNES
d. ST. ÉMILION
30. MONBAZILLAC
31. CAHORS
32. AUVERGNE
33. COGNAC
34. VINS DE LA LOIRE
a. MUSCADET
b. ANJOU
c. VOUVRAY
d. POUILLY
35. CHAMPAGNE
36. BURGUNDY
a. CHABLIS
b. NUITS
c. BEAUNE
d. MONTRACHET
37. BEAUJOLAIS
38. SAVOIE
39. JURA
40. ALSACE
ALEMANIA
41. MOSELA
42. RENANIA
43. FRANCONIA
a. BADEN
SUIZA
44. NEUCHATEL
45. LÉMAN
46. VALAIS
AUSTRIA
47. REGIÓN DANUBIANA
48. BURGENLAND
YUGOSLAVIA
49. ESLOVENIA
50. CROACIA
HUNGRIA
51. TOKAI
URSS
52. MOLDAVIA
53. UCRANIA
PORTUGAL
54. PORTO
55. VINHO VERDE

LAS REGIONES PRODUCTORAS DE VINO de la Europa templada y de la región perimediterránea figuran en este mapa, en el que se distinguen dos líneas de trayectoria casi horizontal, de oeste a este. La línea superior marca el límite septentrional de cultivo de la vid. La línea inferior separa el área en la cual las vides están bajo la influencia del clima mediterráneo de la zona en la que dominan las condiciones de

varse con cierto margen de confianza solamente en un cinturón, de anchura variable, que va desde la línea del límite septentrional hasta otra línea formada por el litoral meridional del Mediterráneo (y, naturalmente, cuando las condiciones son análogas, en otras partes del mundo, en particular California, Chile, Argentina, Africa del Sur y Australia meridional).

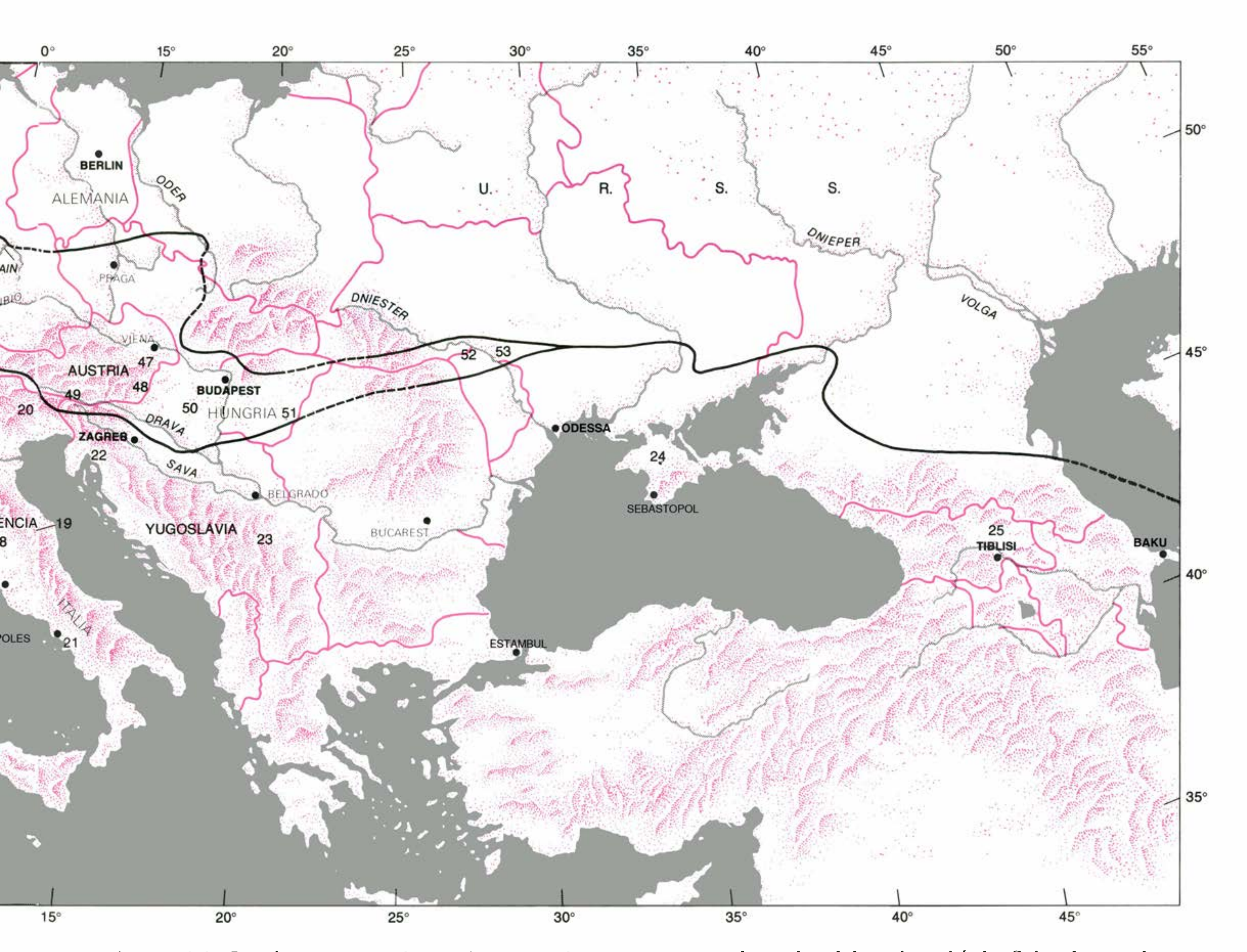
Fijados estos extremos, vamos a trazar otra línea. Esta atraviesa también Europa, de oeste a este. Dicha trayectoria resulta de importancia fundamental para los consumidores que quieren estar orientados al respecto.

La línea divide a Europa en dos regiones vitícolas distintas. Puede tratarse con una precisión considerable (mayor de la que hemos intentado alcanzar en el mapa adjunto), aunque el dibujo aparezca sinuoso o puedan surgir en su recorrido disputas locales a propósito de los puntos de referencia. Paradójicamente, esta línea no se menciona nunca en los libros de consulta habituales, no obstante su importancia fundamental por cómo divide los vinos europeos en dos grupos.

La línea empieza en Portugal, al sur de Oporto, sigue el valle del Duero, región productora del oporto, subiendo hacia el este. Una vez superada la parte central de la Península Ibérica vuelve a girar al norte, cortando la comarca española de la Rioja entre Haro y Logroño. Remontados los Pirineos, reanuda su curso por el país vasco francés encaramándose al noreste en dirección a Cahors, junto al río Lot y cortando el río Garona cerca del punto en donde éste recibe al Tarn, afluente suyo.

Cahors, que tiene relaciones muy antiguas con Burdeos, situado aguas abajo, es un punto axial. A partir de aquí, la línea zigzaguea subiendo y bajando, por la región montañosa del centro, dobla en dirección noreste hasta las cabeceras del Allier y del Loira, bordea el flanco occidental de los montes Cévennes y, finalmente, arriba a Lyon, en la confluencia del Saona (el río del borgoña-beaujolais) con el Ródano, por el norte. De aquí, la línea asciende por el serpenteante valle del Ródano hasta el lago de Ginebra. Posteriormente se borra la trayectoria, que es mejor indicar mediante una línea a trazos cuando atraviesa los Alpes. Los lagos glaciales de la Italia septentrional plantean un problema: probablemente la línea debería pasar al norte de los mismos, pero este punto es discutible.

E, inmediatamente, se encuentra de nuevo el curso de la trayectoria. Avanza definidamente por el norte de la llanura veneciana, penetra en Yugoslavia, cortando la Eslovaquia septentrional por Maribor (cuyo ángulo separado estuvo



clima templado. Los números corresponden a regiones que se destacan en el cultivo de vinos mediterráneos y de vinos de clima templado. Los mismos números figuran en la lista de la izquierda del mapa, en donde aparecen los nombres de las regiones vinícolas. Se introducen en el mapa cierto número de ciudades que sirven de puntos de referencia. Los países son los citados en la lista de la izquierda que acompaña al mapa.

un tiempo sometido a Austria que producía un vino de excepcional calidad), para continuar hacia el este, a lo largo del curso del río Drava. Tras un largo trecho a su vera, por su margen meridional, prosigue por el borde oriental de la llanura húngara, remonta los Cárpatos, cruza Moldavia y acaba uniéndose con la otra línea (el límite septentrional), al norte del mar Negro.

Esta línea define, desde el punto de vista del enólogo, el límite septentrional de la influencia mediterránea. La cuenca mediterránea tiene un clima biestacional con inviernos suaves y lluviosos y veranos cálidos y secos (prácticamente sin lluvia). Tiene una flora característica, con olivos, alcornoques, naranjos, limoneros, plantas de la familia de los melones, adelfas y numerosas plantas de frutos exquisitos. También tiene su propio grupo peculiar de vides, todas ellas con un árbol genealógico que se remonta a los albores de la historia.

Como era de esperar, el clima mediterráneo y el parentesco común de sus vides confieren un intenso parecido de familia a todos los vinos mediterráneos. Por regla general, son suaves, pastosos, de baja acidez, de graduación alcohólica relativamente alta y de poco aroma afrutado. Su aroma presenta un regusto recio y terroso, aunque no desagradable. En su mayoría tienen poco aroma, pero unos pocos (del grupo del muscat) son muy aromáticos. Las diferencias entre las vides de uva negra de las regiones costeras de España, Argelia, Tunicia, Languedoc, Sicilia, la Dalmacia yugoslava y Grecia, por ejemplo, son sólo de detalle. Los vinos son básicamente los mismos. El límite septentrional para los vinos de este tipo viene determinado por la segunda línea que hemos trazado. También en este caso el lector debe tener en cuenta que esta línea no es inamovible y que puede desplazarse unos cuantos kilómetros en una u otra dirección.

Al norte de esta línea, las cosas cambian completamente y hallamos un mundo distinto. Aquí, en una banda estrecha e irregular que se extiende de un confín a otro de Europa, es un clima templado, de cuatro estaciones perfectamente diferenciadas, el que acogerá a las vides; los inviernos son fríos; las temperaturas, muy variables durante el período de crecimiento, y, la lluvia, distribuida en el cuatripartito ciclo anual. Los suelos también son distintos debido a que (amén de las grandes diferencias de origen geológico) sus características tienden a estar en función del clima. El suelo está "embebido en el clima". En esa ancha franja la vida vegetal difiere

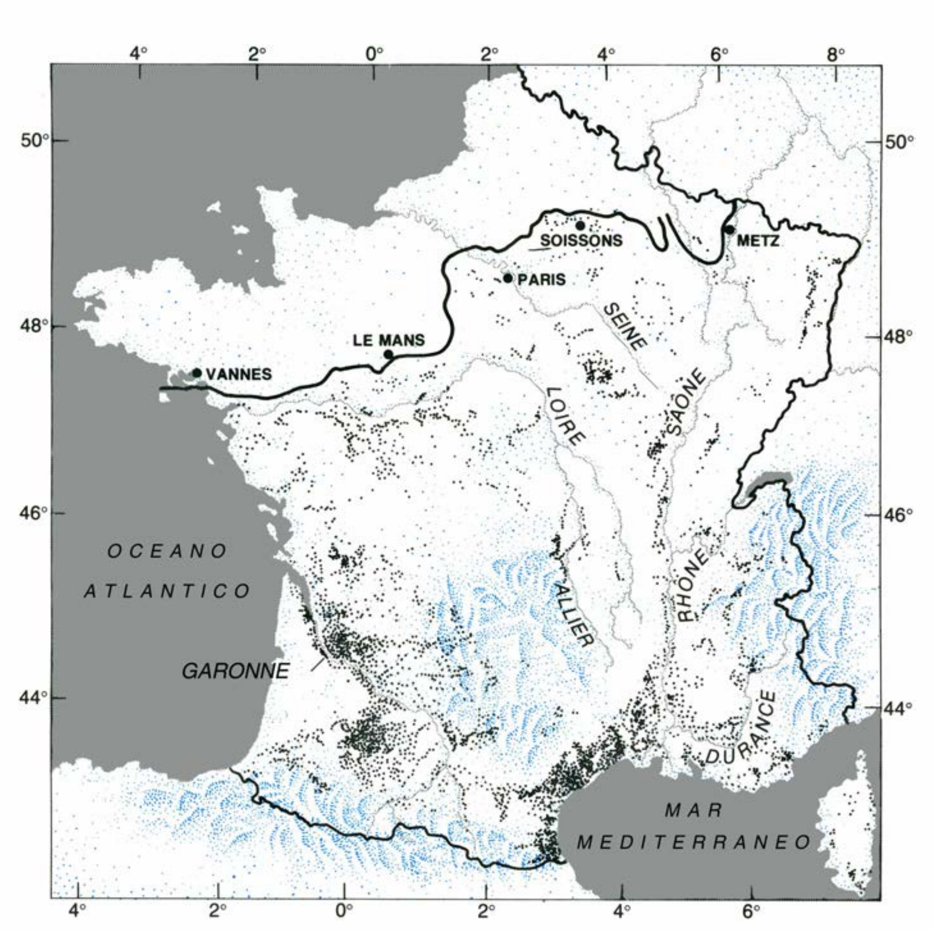


ORIGEN MEDITERRÁNEO		ORIGEN DE CLIMA TEMPLADO	
VID	HECTÁREAS	VID	HECTÁREAS
THOMPSON SEEDLESS	92.370	FRENCH COLOMBARD	9.550
CARIGNANE	11 510	CHENIN BLANC	5.770
ZINFANDEL	9 515	CABERNET SAUVIGNON	4 595
GRENACHE	6 595	PINOT NOIR	1 995
BARBERA	5 895	CHARDONNAY	1.830
MUSCAT	5 015	WHITE RIESLING	1.250
RUBY CABERNET	4 585	GAMAY BEAUJOLAIS	1.100
PETITE SYRAH	3 295	GAMAY	1 020
ALICANTE BOUSCHET	2 760	SEMILLON	995
RUBIRED	2 735	SAUVIGNON BLANC	665
MISSION	2 560	SYLVANER	570
PALOMINO	2.515	GEWURTZTRAMINER	465
SALVADOR	1 070	MERLOT	405

LAS REGIONES VINICOLAS DE CALIFORNIA se designan con sus iniciales que, de arriba abajo, corresponden, respectivamente, a Sacramento Valley, North Coast, Central Valley, San Joaquin Valley, Salinas River Valley, South Coast y Hot Desert. En el recuadro figuran las principales variedades, distinguiéndose entre las que producen vino tinto *(color)* o blanco *(negro)*. California tiene un clima mediterráneo, biestacional, modificado por cuanto el gran valle interior está casi cerrado por montañas; característica que reduce las precipitaciones e intensifica el calor.

también de la que desarrolla la flora mediterránea, sin que a ello se oponga la existencia de una superposición parcial.

Igual ocurre con la uva, que tiene un aspecto diferente, un sabor diferente y produce vinos diferentes. En general, estos vinos son más afrutados, de menor graduación alcohólica y de mayor acidez original que los vinos mediterráneos. Tienen también aromas más sutiles y más delicados. Aunque la cuenca mediterránea tiene sus vinos superiores (los chianti tintos, ciertos vinos del Piamonte y de los lagos glaciales italianos, los mejores vinos del valle del Ródano, y, naturalmente, especialidades tan elaboradas como el sherry y el marsala), su número es relativamente pequeño. Los enólogos coinciden en que la mayoría de los vinos verdaderamente finos (los borgoña, los claretes y champañas, y los vinos de Alemania y Alsacia, por no mencionar otros que son menos familiares, como los de Austria y Hungría) se dan en un clima templado.



ORIGEN MEDITERRÁNEO		ORIGEN DE CLIMA TEMPLADO	
VID	HECTÁREAS	VID	HECTÁREAS
CARIGNANE	182.780	GAMAYS	40.510
ARAMON	158.080	SÉMILLON	29.640
UGNI BLANC	47.425	MALBEC	16.800
GRAND NOIR	39.520	GAMAYS TEINTURIERS	15.810
ALICANTE BOUSCHET	34.580	CABERNET FRANC	13.830
GRENACHE	31.616	FOLLE BLANCHE	11.860
CINSAUT	15.810	SAUVIGNON	10.870
TERRETS	14.820	COLOMBARD	10.870
CLAIRETTES	13.830	CHÉMIN BLANC	10.870
MORRASTEL	10.870	MUSCADET	10.870
MAUZACS	6.420	JURANCON	9.890
VALDIGUIÉ	4.940	FROLLEAU	9.890
SYRAH	3.260	PINOTS	7.900

REGIONES VINICOLAS DE FRANCIA, indicadas por puntos negros en el mapa. El recuadro inferior presenta las variedades más importantes, con el número de hectáreas de viñedo correspondientes a cada variedad y la indicación del color del vino: tinto *(en color)* o blanco *(en negro)*.

La naturaleza general de los vinos producidos a cada lado de la línea está, pues, determinada por una combinación del clima dominante y las vides adaptadas al mismo y, en una medida menor, por el suelo. Las sutiles diferencias en la calidad del vino, dentro de cada una de esas dos grandes divisiones que hemos establecido, se deben a microclimas y microadaptaciones de las uvas. La región de Borgoña nos brinda un buen ejemplo de ello. Los grandes borgoñas, tintos y blancos, se cosechan en la región sudeste, soleada y resguardada, y en los declives orientados al sur de la Côte d'Or, a partir de ciertas subvariedades de unas pinot negra y chardonnay. En el valle del Saona, unos cuantos centenares de metros al este, el microclima cambia, cambian las uvas que se recogen y los vinos no pueden considerarse como grandes caldos. Lo mismo ocurre en la parte superior de las colinas en las pendientes de la vertiente occidental.

Contrariamente a la creencia generalizada, el suelo, aunque importante, sólo es un factor puramente secundario, en comparación con el clima y la variedad de cultivo a que pertenece la vid. Esta relación resulta evidente si tenemos en cuenta que en los viñedos más famosos del mundo se encuentran muchos suelos totalmente distintos unos de otros.

La firmeza de la íntima relación que media entre clima y vid puede ejemplificarse de otro modo. Si la pinot negra, de Borgoña, se trasplanta a otro punto de clima general análogo, verbigracia, la región de Champagne, del Mâeonnais, del Médoc, o del valle del Loira, el vino que se obtendrá podrá presentar un cierto parecido, a pesar de lo cual será netamente diferente. Si esa cepa se planta al otro lado de la línea, en la cuenca mediterránea, la calidad específica de la pinot negra desaparece, y lo que se obtiene no es sino un vino tinto mediterráneo más, sin mayores refinamientos. Algo parecido acontece con otras variedades famosas de uva de clima templado, como la sauvignon blanca, la riesling, la chardonnay y la cabernet. En contrapartida, cepas mediterráneas como la garnacha o el cariñena, al trasplantarse a la demarcación del clima templado, se debilitan y pueden incluso no llegar a madurar por ello.

LOCALIDAD	CLIMA MEDITERRÁNEO											
	EN.	FEB.	MAR.	ABR.	MAYO	JUN.	JUL.	AG.	SEPT.	OCT.	NOV.	DIC.
VALENCIA	**10.4** 30.7	**10.8** 31.1	**13.4** 31.3	**14.5** 33	**18** 28.9	**21.4** 35.4	**24** 8	**24.6** 21.4	**22.6** 49.8	**18** 106.4	**14.2** 57.4	**11.8** 30.9
MARSELLA	**5.7** 38.9	**6.4** 41.3	**10.2** 42.6	**12.7** 33.8	**16.9** 42.7	**20.8** 34.3	**23.3** 15.1	**22.6** 30.2	**19.8** 56.5	**14.8** 80.6	**10.2** 59	**7.9** 72.6
NAPOLES	**6.4** 69	**9.9** 64	**10.3** 38	**13.8** 59	**17.4** 59	**21.4** 21	**23.6** 16	**23.8** 27	**21** 92	**16.9** 118	**13** 129	**9.9** 136
ATENAS	**10.6** 58.8	**11.5** 25.5	**12.3** 32.1	**16.3** 19.1	**20.9** 29	**25.7** 10.8	**28.6** 2.7	**28.9** 6.2	**24.7** 10.6	**19.8** 68.7	**15.6** 64.9	**12.3** 52.4
SALONICA	**6.1** 42	**7.5** 29.8	**9.4** 32.8	**14.3** 34.5	**19.6** 44.7	**24.8** 30	**27.4** 18.5	**27.4** 16.9	**22.5** 30.6	**16.5** 62.9	**11.7** 61.2	**8.2** 49.7
ODESSA	**−4.1** 15.7	**−1.2** 21.2	**2** 22.3	**9.4** 24.1	**16.1** 33	**20.2** 38.6	**22.1** 41.6	**21.4** 42.3	**17.4** 10.1	**10.4** 28.6	**4.8** 37.4	**−.2** 35
SACRAMENTO (CALIF.)	**8.4** 112.7	**10.4** 78.1	**12.4** 52.6	**15.5** 48.5	**18.4** 14.9	**21.8** 3.3	**24.4** T	**23.4** 2.7	**22.5** 12.8	**18.1** 13.3	**12.2** 41.7	**8.7** 91.7
BAKERSFIELD (CALIF.)	**8.5** 29	**11.1** 29	**13.8** 27	**17.3** 21	**21.3** 6	**25** 2	**29** T	**27.7** T	**24.7** 2	**19.3** 8	**13.3** 12	**9.3** 25
NAPA (CALIF.)	**8.6** 123.4	**10.3** 108.4	**11.9** 84.3	**13.6** 43.9	**16** 22.1	**18.3** 5.3	**19.5** .25	**19.1** .76	**19.2** 4.6	**16.5** 30.9	**12.5** 57.6	**9.4** 124.7
YAKIMA (WASH.)	**−2.6** 30.2	**1.1** 22.1	**5.5** 15.7	**10.3** 11.9	**14.7** 13.7	**18** 20.3	**21.6** 3.3	**21.1** 5	**16.3** 8.9	**10.3** 15.2	**3** 24.4	**−2.7** 28.4

LOCALIDAD	CLIMA TEMPLADO											
	EN.	FEB.	MAR.	ABR.	MAYO	JUN.	JUL.	AG.	SEPT.	OCT.	NOV.	DIC.
REIMS	**1.6** 45.7	**3.3** 38.1	**7.7** 48.3	**10** 48.3	**13.3** 53.3	**16.6** 53.3	**18.8** 66	**18.3** 58.4	**15.5** 43.2	**10** 66	**5.5** 58.4	**2.2** 58.4
DIJON	**1.5** 61.1	**2.2** 44.8	**7.1** 50	**10.2** 36	**14.6** 54.8	**17.5** 97.6	**19.6** 50.8	**18.6** 82.9	**15.8** 75.9	**10.4** 62.5	**5.6** 63	**3.8** 65.9
NANTES	**4.9** 80.5	**5** 73.7	**8.5** 58.4	**10.5** 47.4	**14** 46.8	**16.9** 50.6	**18.5** 52.8	**18.1** 73	**16.1** 88.1	**12.3** 74.9	**7.9** 86.4	**6.8** 108.5
BURDEOS	**5.2** 101.3	**5.4** 88.1	**9.5** 67	**11.3** 41.3	**14.8** 63.8	**17.6** 76.6	**19.1** 69.6	**19** 75.3	**17** 87.7	**12.6** 86.3	**8.1** 85.2	**6.8** 125.5
FRANKFURT	**−.3** 53.4	**1.9** 50.6	**6.2** 35.1	**11.3** 39	**14.9** 58.5	**17.7** 76.9	**19.7** 67.7	**19** 74.7	**15.7** 43.9	**10.4** 30.9	**5.1** 62.4	**1.5** 58.8
GINEBRA	**1.1** 63	**2.2** 56	**6.1** 55	**10** 51	**14.1** 67	**17.8** 89	**19.9** 64	**19.1** 94	**15.8** 99	**10.3** 72	**5.7** 83	**2.1** 59
VIENA	**−.4** 34.6	**.1** 34.2	**4.4** 46.6	**9.9** 50.7	**14.4** 55.9	**18** 83.5	**19.9** 87.8	**19.2** 72.4	**15.3** 33.9	**10** 41.6	**4.9** 40.2	**2** 48.7
BUDAPEST	**−.2** 41.3	**.6** 38	**5.5** 33.5	**11.4** 39.5	**16.2** 61	**20.2** 94.1	**22** 66.4	**21.3** 56	**17** 39.2	**11.1** 49.8	**5.8** 57	**2.6** 54.4
GENEVA (N.Y.)	**−4.1** 50.2	**−2.8** 63.6	**.3** 77.6	**8.4** 69.8	**13.4** 71	**19.3** 78.4	**21.8** 78.2	**20.8** 69.9	**16.9** 63.1	**11** 89.2	**4.8** 51.7	**1.6** 60
BALTIMORE	**1.5** 86.3	**2** 73.6	**6.1** 96.5	**12.3** 91.4	**18** 101.6	**22.5** 83.8	**24.8** 106.7	**23.8** 132.1	**20** 83.8	**13.8** 81.2	**7.5** 78.7	**2.1** 76.2
SANDUSKY	**−2.1** 68.9	**−.7** 57.9	**2.5** 73	**10.2** 86.1	**15.6** 79.7	**21.5** 88	**23.8** 87.2	**23.1** 97.6	**19.2** 50.9	**13.1** 52	**5.6** 59.4	**0** 56.1
CHARLESTON (S.C.)	**9.4** 73.9	**11.1** 88.3	**13.2** 108	**18** 72.7	**22.4** 97.7	**25.3** 130.5	**26.7** 199.8	**26.5** 134.2	**23.9** 164.9	**18.6** 74.5	**13.3** 43.7	**9.7** 75.7
ST. LOUIS	**−.2** 36.4	**2.5** 64.7	**5.5** 66.9	**13.6** 86.4	**19.1** 83.8	**24.5** 111.6	**26.6** 95.9	**26** 56.5	**21.7** 46.8	**14.8** 63	**6.8** 59.8	**2.1** 42.1
LUBBOCK	**4** 17	**6.2** 15	**9.8** 18	**15.2** 29	**20.1** 81	**25.3** 64	**26.4** 51	**26** 43	**21.9** 60	**16.2** 51	**8.8** 14	**5** 16
ALBUQUERQUE	**1.4** 8.6	**4.8** 10.9	**7.6** 12.2	**13.3** 16.5	**18.4** 20.8	**23.2** 11.7	**25.6** 34.2	**24.7** 29.8	**21.1** 21.8	**14.6** 17.8	**7.1** 5.6	**2.7** 13.2
BISMARCK	**−13.6** 14.3	**−9.6** 12.2	**−4.7** 16.2	**6** 21.5	**13** 60.6	**17.8** 86	**21.8** 48.7	**21.1** 51.5	**14.4** 25.5	**7.9** 15.3	**−2.1** 18.6	**−7.7** 10.7
BOISE	**−.9** 40.4	**1.4** 32.2	**4.5** 28.6	**9.4** 31.3	**13.9** 43.1	**18.2** 22.9	**23.8** 6.3	**21.9** 6.3	**17.6** 8.5	**11.1** 17.5	**3.6** 26	**−.1** 33.2
PORTLAND (ORE.)	**4.9** 211.8	**6.6** 118.1	**7.8** 124.7	**11.6** 66.9	**14.8** 52.9	**17.1** 47.9	**20.2** 6.8	**19.7** 22.6	**18.2** 42.2	**13.8** 97.7	**8.3** 142.9	**6.1** 169.5

CONDICIONES CLIMATICAS predominantes en regiones vinícolas típicas, indicadas por los datos de temperatura media *(negrita)*, en grados centígrados y precipitación media *(redonda)*, en milímetros. La *T* que aparece a veces en estos últimos significa precipitación inapreciable.

Esta adaptación de las cepas justas a un clima o microclima dados constituye el objeto de estudio de la ampelografía, que pertenece ante todo a la botánica para continuar luego relacionada con la climatología, la geología y la historia humana. Permítasenos analizar algunos de estos factores.

Retrocediendo hasta los documentos fósiles, se han encontrado restos de *Vitis* (de especies actualmente extinguidas) en casi toda Europa, incluida la misma Escandinavia. Tras un largo y prolongado proceso de migración y adaptación, otros dos grupos de especies relacionadas con las europeas se desarrollaron en Extremo Oriente y en América del Norte. La relación entre ellas es indudable, porque todas las especies del género *Vitis,* con una sola excepción, tienen idéntico número de cromosomas (38) y son interfecundas. (Las vides muscadine del sudeste de Estados Unidos constituyen la excepción: poseen 40 cromosomas y no hibridan con las demás.) Los grupos del Asia oriental y de América del Norte siguieron un curso evolutivo divergente, y en su forma pura carecen de interés para la vinificación.

La evolución de las estirpes eurasiáticas se vio drásticamente interrumpida por los cambios bruscos que tuvieron lugar en el clima con los períodos glaciales. Cada glaciación provocó una retirada de la vid hacia el sur, hasta una u otra de las zonas de refugio, la primera en la cuenca mediterránea y la segunda al sur del mar Caspio. Cada vez que los hielos se retiraban hacia el Norte ocurría una nueva expansión de cepas salvajes, en formas modificadas. A partir de ellas acabó diferenciándose una nueva especie (*V. sylvestris,* que se clasifica también como *V. vinifera* ssp. *sylvestris*), que, a su vez fue adaptándose a las numerosas variaciones climáticas, dando así origen a un cierto número de grupos diferentes. Se ha podido saber, basándose en el estudio de hallazgos de semillas, que el fruto de *V. sylvestris* se recolectaba y consumía por poblaciones prehistóricas.

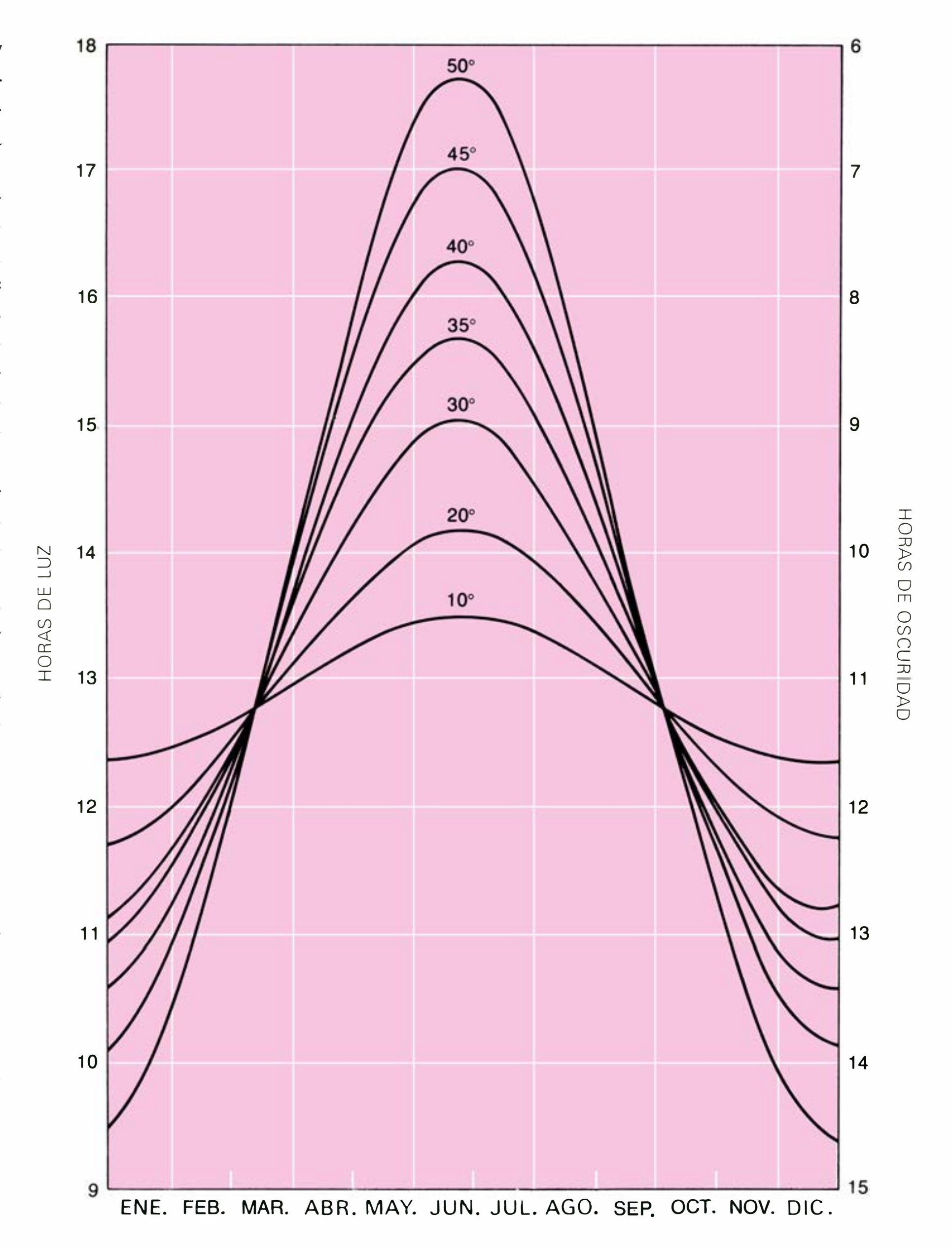
Las vides salvajes (con frecuencia denominadas *lambruscas,* corrupción del latín *labrusca,* cepa salvaje) continuaron medrando hasta bien entrada la mitad del siglo XIX (hay muchas cepas salvajes distribuidas en la actualidad por la mayor parte de América del Norte). Pero a partir de entonces quedaron prácticamente eliminadas por la filoxera, un insecto introducido de forma accidental desde América, y frente al cual carecían de defensa. Incluso en nuestros días pueden encontrarse ejemplares supervivien-

tes en algunas regiones de los Balcanes y de los Pirineos y en otros puntos aislados. Encontrar una equivale casi a tropezar con una tribu neolítica hasta ahora desconocida.

Las vides que permanecieron rezagadas en las zonas de refugio evolucionaron de forma totalmente distinta. Eran sensibles a los rigores del invierno y, de acuerdo con la latitud, maduraban los frutos en días de escasa luz solar. Presentaban un follaje liso, no coriáceo, un sarmiento macizo y arborescente, no trepador, y exigían una temporada de crecimiento y maduración más prolongada.

Se produjo entonces un cambio brusco e importante. En su forma salvaje las vides son dioicas, es decir, existen plantas masculinas y plantas femeninas, y, para producir fruto, la planta funcionalmente femenina debe ser polinizada por una vid masculina, provista de estambres. Sin embargo, hace varios miles de años, en el refugio del Caspio, aparecieron uno o más mutantes de la cepa salvaje originando cepas monoicas: las flores reunían el pistilo, u ovario, de la vid funcionalmente femenina y los estambres, formadores de polen de la vid masculina; con lo cual, las plantas estában en condiciones de autofecundarse (la historia de este desarrollo ha sido ampliamente reconstruida, pieza a pieza, por el ampelógrafo ruso A. M. Negrul y sus colaboradores).

Esta innovación evolutiva revistió una importancia crucial, por varias razones. Las cepas pueden multiplicarse vegetativamente (sin reproducción sexual), es decir, un esqueje leñoso echará raíces y la vid que así se origine reproducirá exactamente la vid progenitora. Por este camino se puede multiplicar, uniforme e indefinidamente, cualquier mutante o híbrido cuyas propiedades sean interesantes. Así se pudo seleccionar, cultivar y multiplicar una variedad monoica especialmente favorable y conseguir todas las ventajas de la uniformidad y de una autopolinización regular y fiable. Probablemente, tales serían los comienzos de la viticultura.



LONGITUD DEL DIA en diferentes latitudes del hemisferio boreal. Las latitudes clave para la vid son 30 grados, que corresponden al límite meridional para los vinos mediterráneos, y, 50 grados, latitud aproximada de Maguncia y de Praga, que representan el límite septentrional para los vinos de clima templado. Las vides que se adaptan a una región situada entorno a los 30 grados son plantas de día corto, en tanto que las de día largo viven mejor en latitudes superiores.

Una segunda ventaja del hermafroditismo estriba en que puede introducirse por polinización cruzada normal en el patrimonio genético de casi todas las cepas salvajes que carezcan de él, pero posean otras características deseables, tales: dar vino de buena calidad o fruto comestible. El camino quedaba expedito. La capacidad de autopolinización se fue extendiendo a saltos (no uniformemente) en las poblaciones salvajes dioicas, que continuaban existiendo. El hombre se limitó a ir seleccionando. Cuando aparecía una nueva variedad interesante, era cosa fácil extenderla por multiplicación vegetativa.

La selección dependió de las preferencias culturales. En ciertas culturas se puso el énfasis en la obtención de grandes racimos para el consumo directo o para el secado y conversión en pasas; en otras, el interés se centró en la calidad del vino. Negrul y colaboradores han demostrado que las plantas de granos voluminosos, entre las que se cuentan algunas cuyos granos carecían de semillas, estaban concentradas en una zona situada al sur del mar Caspio. A este grupo de estirpes se les reconoce comúnmente hoy por *Prol.* (del latín *proles,* que significa progenie) *orientalis.* Los tipos de grano más pequeño, más adecuados para la vinificación, se concentraron en la cuenca del mar Negro; Negrul dio a este grupo el nombre de *Prol. pontica.*

Las vías evolutivas se reconstruyeron (si bien todavía de un modo incompleto) principalmente sobre la base de estudios muy detallados de las semillas de uva, que en ampelografía constituyen el equivalente de los dientes en paleontología humana, puesto que también son estructuras persistentes. Los restos arqueológicos demuestran que la recolección de uva salvaje continuó incluso después de la explotación de las formas cultivadas. Se han encontrado asimismo

vestigios de una industria vinícola especializada, que en el siglo VI a. C. se había convertido en la industria principal de muchos pueblos, situados especialmente en torno al mar Negro. Se han excavado restos de todo cuanto dicha industria necesitaba en las localidades de Quersoneso (a cinco kilómetros al oeste de Sebastopol), Mirmekion y Tiritica. Se han descubierto cuchillos y prensas para las uvas, depósitos de fermentación y almacenaje, de piedra, arcilla y cerámica, así como plataformas en donde se pisaba la uva. Algunos de ellos se parecen al equipo que todavía se emplea en determinadas regiones vinícolas, y un viñador experto podría reconocer inmediatamente cualquiera de estos artefactos.

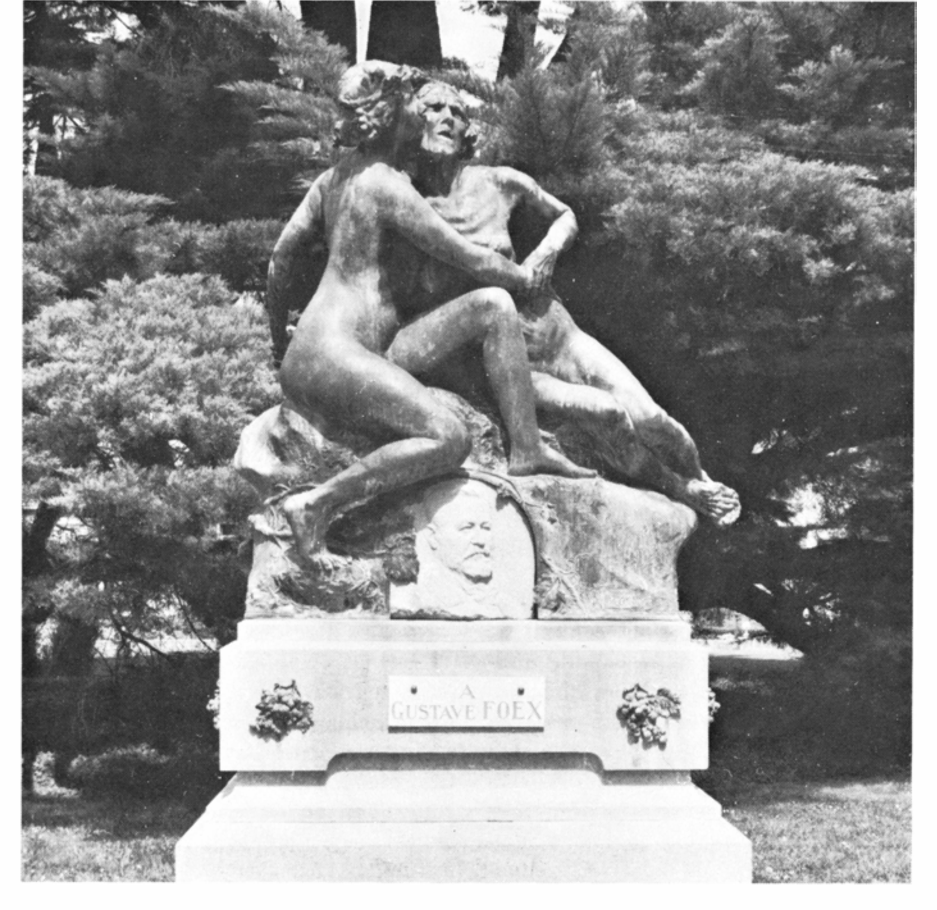
Desde esta patria de la viticultura, se observó una migración gradual en dirección al oeste, en parte de forma natural y en parte con la colaboración de personas que llevaban consigo vástagos de vid. A través de una continua polinización cruzada entre las vides orientales y las vides salvajes de *V. sylvestris* encontradas más a occidente a lo largo de la vía seguida, la cuenca mediterránea se pobló de vides del tipo predominante hoy todavía. La Península Ibérica fue especialmente rica en estos híbridos tan útiles, muchos de los cuales iniciaron una migración inversa hacia oriente, a lo largo de dicha cuenca. Ejemplos notables entre estos tipos más occidentales (la *Prol. occidentalis*) son los garnacha (y otros muchos sinónimos), los cariñena (y otros nombres) y los aramón, que siguen produciendo torrentes de vino, bastante común, en muchos puntos de la cuenca mediterránea.

La evolución de la *V. sylvestris* salvaje en la Europa templada dando origen a distintos grupos ecológicos cubre otro capítulo de esta historia. Las vides del noreste de Francia fueron diferenciándose de las del sudoeste, y así sucesivamente. Al penetrar en la Europa templada, los pueblos mediterráneos llevaron consigo sus vides, como era de esperar. En su mayor parte estas cepas resultaron inadaptadas al nuevo clima, pero pudieron resistir el tiempo suficiente para cruzarse con las cepas salvajes indígenas. De la conjunción resultante, aparecieron estirpes salvajes que combinaban cualidades interesantes de cada progenitor: autopolinización, capacidad de maduración del fruto en días con muchas horas de luz solar, bien adaptadas a un clima templado, dotadas del vigor de los híbridos y provistas de un surtido de nuevos aromas y sabores de especial valor para el vino.

A partir de estas vides, se seleccionaron las variedades cultivadas en la Europa templada. Los nombres y sinónimos de muchas plantas de éstas dan fe de su origen salvaje: plant des bois, sauvaget, sauvignon, savagnin, sauvagnou, fer servadou y lambrusquet. Un filólogo que conozca bien los dialectos europeos podría agregar muchos más.

Otra prueba de este origen casual se encuentra en la inmensa variedad de lo que los franceses denominan un *cépage,* que no corresponde a una variedad única y uniforme que proceda de una primera cepa, sino a un grupo de individuos parecidos aunque no absolutamente idénticos. Mencionemos, como ejemplo, la variedad pinot tinto. Este nombre no se aplica a un individuo concreto sino a un conjunto de ellos que tienen en común un mismo fondo genético. Algunos de ellos se denominan pinot con una localidad adjunta (pinot D'Aunis, pinot d'Ambonnay). Los nombres de otros pinot añaden al nombre del lugar otros caracteres especiales (Gros Plant doré d'Ay) o el nombre del que popularizó la variedad (Pinot Renault). En la región de Champague se encuentran docenas de pinot, otras docenas en la Côte d'Or y no menos de cuarenta sólo en la región de habla alemana de Suiza. En total, fácilmente llegan a mil los vinos que tienen derecho al nombre de pinot tinto. Así ocurre con los nombres de otros *cépages* famosos: los generalizadísimos chasselas, los côtes, los numerosos carmenet (incluido el cabernet), los riesling. Cada *cépage* comporta una multitud de primos, hermanos y tíos; cada variedad hunde sus raíces en una cepa o en una mutación que progresó desapercibida o llamó la atención a alguien alguna vez, y quedó fijada inmediatamente. Todo ello parece muy complicado, pero así es en realidad. La comprensión real y la justa apreciación de las viñas, con sus diferencias espectaculares o sutiles, empiezan con un hondo entendimiento de estos reajustes inacabables entre la vid y el clima –y, en menor grado, entre la vid y el suelo– y con un conocimiento de los hombres que cultivaron y desvelaron sus secretos de generación en generación.

Hasta aquí, la historia no ha sido otra que la de la selección natural, por la que



CONJUNTO ALEGORICO que se halla en la escuela de viticultura de Montpellier. Representa una mujer joven que sostiene a otra, anciana. La joven simboliza las cepas americanas que revitalizaron las francesas, representadas por la anciana, tras la devastación de éstas por la filoxera en el siglo XIX. Gustavo Foex fue un experto viticultor que dirigía la escuela en aquel entonces.

el hombre se limitaba a escoger entre los resultados que se presentaban espontáneamente. La historia ha tenido que apoyarse, en varios de sus puntos, en conjeturas e inferencias. Conviene detenerse en la intervención deliberada del hombre en el proceso genético, ocurrida en el siglo pasado. Mediante cruzamientos intraespecíficos en la especie eurasiática e hibridaciones entre especies distintas se crearon variedades totalmente nuevas.

La devastación debida a la epidemia de la filoxera, que arrasó las cepas europeas a finales del siglo XIX, señala el verdadero punto de partida. Se manifestó entonces la necesidad desesperada de encontrar rizomas resistentes a la filoxera, en los cuales pudieran injertarse las variedades clásicas, lo que permitiría así su supervivencia. Ello implicaba recurrir a las cepas americanas, no sometidas todavía a procesos de selección, pero que eran, y son, resistentes en diverso grado a la filoxera, toda vez que se habían adaptado a dicho insecto, que es autóctono del Nuevo Mundo.

Encontrar rizomas que fueran compatibles, y por tanto adaptables, con las condiciones europeas exigía combinaciones y recombinaciones. El hecho de combinar un carácter de *V. rupestris,* de los Apalaches, por ejemplo, con otro carácter de *V. berlandieri,* de Texas, y cruzar la planta descendiente con otras variedades hasta la quinta o sexta generación, se convirtió en algo rutinario. Resultado de este trabajo fueron los rizomas reconocidos por nombres crípticos (A × R 1, E.M. 41, Rupestris St. George, C.3309, 5BB, etc.) gracias a las cuales los viticultores de Europa, California, Australia, Sudáfrica y Chile medran actualmente.

Paralelamente a esta labor, se realizó un esfuerzo divergente y competitivo, que consistió en superar la necesidad de realizar injertos, mediante la producción de híbridos que combinaran la calidad del fruto y del vino de las vides europeas, con la resistencia a los rigores invernales y a las enfermedades de estas otras especies foráneas. Como es fácil de comprender, estos híbridos no solamente extenderían el cultivo de las vides por nuevas regiones, sino que ofrecerían también ventajas sobre algunas variedades que existían desde antiguo.

El esfuerzo realizado en ambos frentes se vio hostigado. Los viñadores establecidos no deseaban la competencia de nuevas regiones. En términos de calidad, los productores de las *grandes denominaciones de origen* estaban harto satisfechos con lo que tenían, ahora que sus vides podían sobrevivir a base de las nuevas cepas resistentes.

Surgieron otras complicaciones: la hibridación dio, al principio, resultados más bien mediocres, remedando la advertencia de George Bernard Shaw a Ellen Terry, de que un hijo de ambos podía heredar la inteligencia de ella y la belleza de él, y no el resultado contrario según deseaba la actriz. La firmeza y la resistencia a las enfermedades mejoró, pero a costa de la calidad. Sin embargo, los trabajos continuaron, principalmente por obra de un pequeño grupo apasionado y como poseído de un extraño fervor.

En tales cruzamientos cada nueva generación mejoraba los resultados de la anterior. Hay en la actualidad numerosísimos híbridos de este tipo, cuya genealogía suele ser muy complicada, que producen buen vino en condiciones que resultan hostiles a la *V. vinifera.* No se trata de grandes vinos; hasta ahora ningún híbrido produce un vino comparable a los grandes borgoñas y chiaretti o al vino del Rhin. Una de las causas estriba en que algunas uniones genéticas introducen, a la vez, caracteres buenos y malos. Sin embargo, vale la pena recordar que la mayoría de los vinos del mundo proceden de vides de la especie *V. vinifera* y que el 90 % de ellos están lejos de poderse considerar grandes vinos. Existe la tendencia a comparar los híbridos peores con las mejores vides de *V. vinifera,* y no lo contrario, como sería de desear.

Los viticultores europeos de mayor arraigo se han mostrado siempre implacablemente hostiles a este esfuerzo. En las regiones vinícolas de Europa sólo se permite el cultivo de contados híbridos, boicoteándose con prohibiciones y limitaciones la continuación del trabajo privado en el campo de la hibridización, lo que lo hace prácticamente imposible. Esta actitud resulta extraña si se presta atención a las grandes aportaciones que la genética aplicada ha brindado a otras ramas de la agricultura y de la ganadería.

No hay ninguna duda de que la Europa viticultora puede pasarse sin los híbridos. En otras regiones, en especial en Estados Unidos, la hibridación se está realizando intensamente. En California se está trabajando, sobre todo, en la producción de nuevos *métis* (mestizos o cruces intraespecíficos) hibridando vides de la Europa templada y del Mediterráneo, a fin de obtener mejores cepas adaptadas a la región fértil, aunque subdesértica, del Central Valley. Se han obtenido algunos resultados notables. Harold P. Olmo, de la Universidad de Davis, California, y Elmer Snyder, del Ministerio de Agricultura de los Estados Unidos, dirigieron el proyecto. Incluso en regiones de la costa occidental, en donde la vid eurasiática está magníficamente adaptada, la exigencia de resistencia a la filoxera y a los nemátodos inspira un trabajo continuo de hibridación intraespecífica e interespecífica.

Además, en particular en la estación experimental del Ministerio de Agricultura, en Geneva, Nueva York, la investigación enológica se concentra en los cruzamientos interespecíficos, que es la hibridifación en sentido estricto. La razón de ello estriba en que la vid eurasiática no se encuentra realmente adaptada al nuevo ambiente estadounidense, con excepción de unas pocas zonas, sobre todo en California, en donde se encuentran muchas variantes del clima mediterráneo. Los extremos climáticos de tipo continental que prevalecen en la mayoría de otras partes del continente se oponen al desarrollo y maduración de las vides.

Los experimentos con híbridos, como muy bien puede suponerse, no se han reducido al área geográfica norteamericana. Además, desde el punto de vista de la más estricta genética de poblaciones, la hibridación aporta una serie importante de valores a la constitución génica de la cepa en cuestión, según confirman trabajos de laboratorio recientes. Los propios viticultores saben que los híbridos presentan a menudo un vigor superior a las variedades paternas; vigor cuyas causas profundas todavía se desconocen.

Hasta la introducción de los mejores híbridos franco-norteamericanos, dotados de mayor firmeza y resistencia a las enfermedades no se pudo producir en la parte templada de América del Norte con el éxito que razonablemente cabía esperar, vinos parangonables a los niveles de calidad. Hoy se consiguen éxitos en muchas zonas estadounidenses, a veces realmente inverosímiles: en los altiplanos de Texas, en Utah y en los Ozarks, y en casi todo el territorio al este del río Mississippi, New Hampshire y Vermont incluidos. A nadie pueden satisfacerle plenamente los vinos que allí se elaboran, porque el híbrido "ideal" todavía no existe. Ahora bien, el consumidor exigente puede esperar caldos cada vez mejores, ya que se siguen reproduciendo los mejores híbridos actuales, entran en explotación nuevas regiones viticultoras y existe ya un mercado para lo que producen.

España es un país que produce gran cantidad de vinos, de múltiples variedades. Variedades que obedecen a un clima muy divergente según las regiones y a una geomorfología del suelo también dispar, sin descontar la distinta forma de elaboración de los caldos.

En la cornisa cantábrica, se cultiva aproximadamente el 0,5 por 100 del total español de uva, tanto de mesa como de transformación. La gradación alcohólica oscila entre los 7-10°, siendo su volumen de producción en 1974 (últimas cifras oficiales) de 233.608 hectolitros. Las cuatro provincias gallegas representan con sus 31.454 Ha el 2,25 % de la superficie nacional; produce chacolí y vinos tintos, con una gradación de 7 a 10°, siendo su cosecha en 1974 de 1.897.254 Hl.

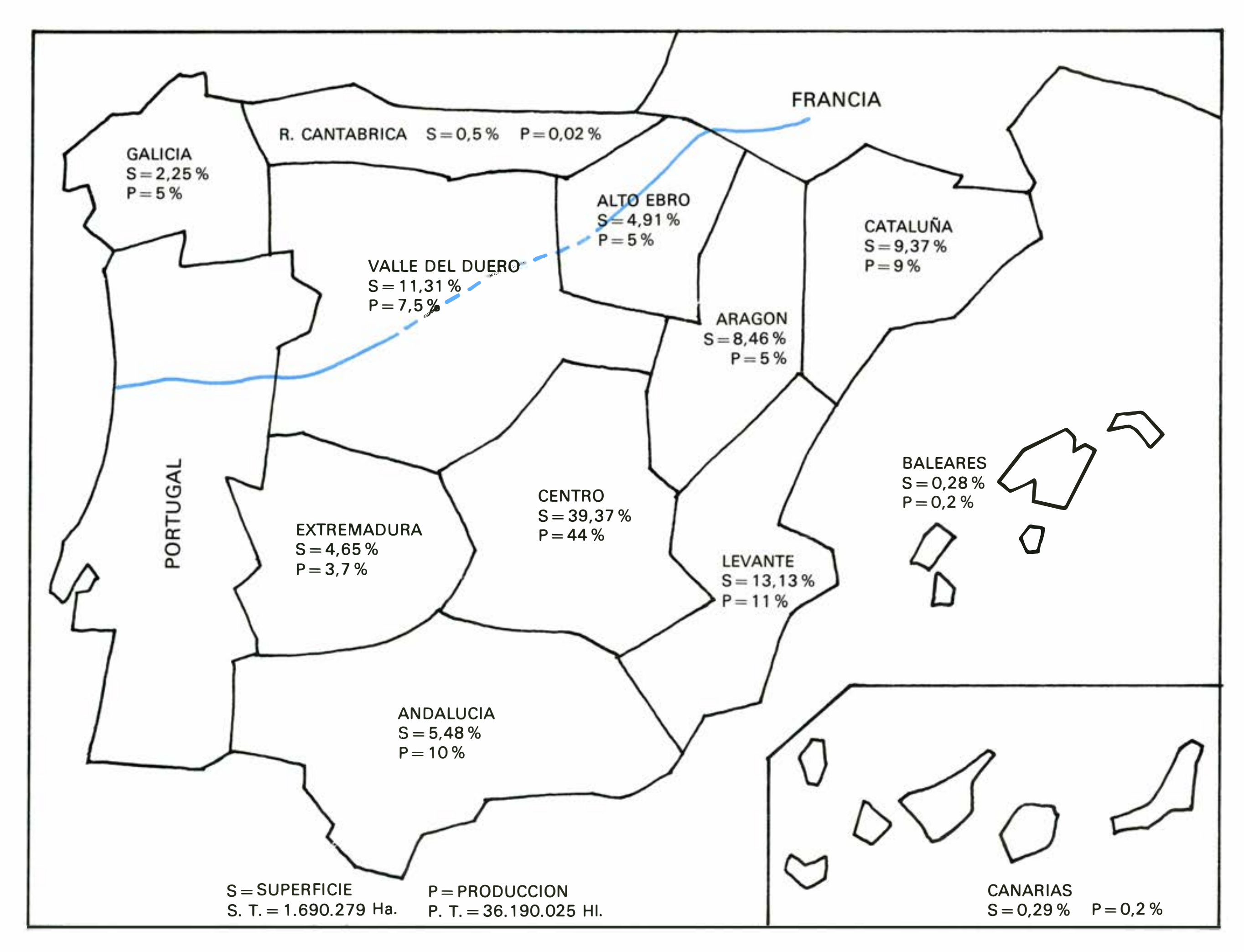
La zona del Ebro, que incluye Aragón y el Alto Ebro, representa un 13,37 por 100 (183.908 Ha) del cultivo de vid. Produce preferentemente vinos tintos, claretes y rosados. La gradación varía entre 10-15°, y su volumen de producción en 1974, fue de 3.664.437 Hl. El nordeste, que engloba las cuatro provincias catalanas y las islas Baleares, cultiva una superficie total de 117.388 Ha (9,65 % del total nacional) con un volumen cosechado en 1974 de 3.673.580 hectolitros; produce vinos espumosos y gasificados. La gradación de sus caldos varía entre los 9 y los 18°.

Las provincias de Avila, Burgos, León, Palencia, Salamanca, Segovia, Soria, Valladolid y Zamora componen la cuenca del Duero. Sus viñedos se extienden por una superficie de 154.502 Ha, un 11,31 del total de España. La gradación de sus vinos varía de 7 a 13°. Su producción en el año de referencia alcanzó los 2.157.501 Hl. Predominan los vinos tintos, rosados y claretes.

La zona Centro, con sus 715.621 Ha de viñedo, acapara el 39,37 por 100 aproximadamente del total. Abarca las provincias de Albacete, Ciudad Real, Cuenca, Guadalajara, Madrid y Toledo. En 1974 produjo 14.965.135 Hl, con una gradación alcohólica que oscila entre los 12 y los 18°. Abunda el vino blanco común.

La región levantina, con las tres provincias valencianas y Murcia, explota 255.350 Ha (13,13 %), con un rendimiento total de 4.158.481 Hl en el año muestra; predominan los tintos, claretes y rosados; su gradación oscila entre 12 y 15°. Las dos provincias extremeñas cultivan el 4,65 por 100 (98.576 Ha) del total español; su producción en 1974 fue de 1.671.600 Hl. Sólo produce vino blanco común, con gradaciones entre 12 y 15°.

Andalucía explota 114.689 Ha, siendo el volumen de producción de vino nuevo en 1974 de 3.697.474 Hl, casi todo blanco. Predominan los vinos generosos secos y dulces, con una gradación entre 12 y 15°; mediante prácticas de elaboración, algunos llegan a alcanzar los 20°. La producción isleña canaria se reduce al 0,29 por ciento (10.572 Ha), con un volumen de 70.955 Hl, blanco común en su mayoría y de una gradación de 10 a 13°.



EL VINO ESPAÑOL se desarrolla en un terreno pobre, árido o semiárido, calizo y arenoso, poco fértil. La vid se cultiva en todas las regiones, cualquiera que sea su latitud. Al ser España un país de microclimas, las cepas se han adaptado a una pluviosidad que supera los 1000 mm en unos casos o no llega a los 300 mm en otros, a inviernos crudos y a veranos tórridos. Las condiciones de suelo y clima son causa de la diversidad de la producción vitícola española. La línea de trazo grueso divide la Península en vinos de clima mediterráneo y de clima templado.

Juegos matemáticos

La base combinatoria del "I Ching", el libro chino de la adivinación y la sabiduría

Martin Gardner

El *I Ching* (se pronuncia *ye ying*) o *Libro de los cambios,* es uno de los libros más viejos del mundo y, también, uno de los más enigmáticos. Durante más de 2000 años se ha utilizado en el Oriente como libro de adivinación, y todavía se estudia con gran respeto como fuente rica en sabiduría confuciana y taoísta. Decenas de miles de jóvenes en Estados Unidos (particularmente en California), que secundan el renacimiento actual del ocultismo y se muestran ansiosos por conocer más sobre el misticismo oriental y sobre la historia antigua de China, consultan hoy el *I Ching* con la misma seriedad que consultan la tabla Ouija o las cartas de tarot. C. G. Jung estaba convencido del extraordinario poder de *I Ching* para predecir el futuro; el mismo lo utilizaría para una prospección sobre la venta en América de una nueva traducción inglesa del libro y obtuvo una respuesta optimista. Eruditos más próximos en la historia que han profundizado en el ocultismo –Colin Wilson, por ejemplo– nos han transmitido sus propias experiencias sobre la clamorosa y profética precisión del *I Ching*.

Se desconocen los orígenes *del I Ching.* Es muy probable que empezara, hacia el siglo VIII antes de Cristo, como una colección de textos, de signo pronóstico, alumbrada en un ambiente agrícola; lentamente, a lo largo de los siglos, dichos documentos se fueron combinando con prácticas adivinatorias realizadas por medio de palillos. Pocos siglos antes de Cristo, hacia el final de la dinastía Chou, adquirió su actual forma, convirtiéndose en uno de los grandes clásicos del canon confuciano.

La base combinatoria del *I Ching* consta de 64 hexagramas que muestran todas las permutaciones posibles de dos tipos de líneas, al tomarlas de seis en seis. Cada hexagrama tiene un nombre chino tradicional. Los dos tipos de línea revelan la dualidad básica de la metafísica china: la línea discreta o cortada corresponde al yin y la línea continua al yang. Tomando las líneas de dos en dos, hay $2^2 = 4$ maneras de combinarlas, a las que se llama digramas, y $2^3 = 8$ maneras de formar trigramas. Los trigramas, con sus nombres chinos y sus significados simbólicos, aparecen en la ilustración de la página 113.

Existen dos maneras antiguas de presentar los ocho trigramas en un círculo. La más antigua, conocida como la colocación Fu Hsi, en memoria del mítico fundador de la primera dinastía china (la dinastía Hsia 2205-1766 a. C.), es, de las dos que figuran en la parte inferior de la página la de la izquierda, y, también, la de la cubierta de la revista. Nótese que los pares opuestos son complementarios, en su significado simbólico y en sentido matemático, pues cada uno se obtiene del otro por substitución de las líneas yin con las yang, y las yang con las yin. Esta disposición, que circunda normalmente el símbolo familiar del yin-yang, se sigue utilizando ampliamente en toda China, Japón y Corea, como un amuleto de buena fortuna, para colocarlo sobre los portales y en joyería. Se le llama también el "cielo más antiguo" o la colocación primaria. La colocación Rey Wen (en memoria del padre legendario, fundador de la dinastía Chou), que aparece a la derecha, en la ilustración inferior de la página 114 (llamada también el "cielo más reciente" o "colocación del mundo íntimo"), abandona la colocación complementaria de la secuencia Fu Hsi, de modo que los trigramas de los puntos cardinales de la brújula simbolizan las estaciones en orden cíclico. Empezando por el sur (tradicionalmente colocado arriba), y en el sentido de las agujas del reloj, los hexagramas que aparecen en los puntos cardinales representan el verano, el otoño, el invierno y la primavera.

La manera más antigua de disponer los 64 hexagramas, conocida como la secuencia Rey Wen, es el orden en el que aparecen en el *I Ching (véase la página siguiente).* Las filas se toman de derecha a izquierda, según indica la numeración. Obsérvese que los hexagramas están apareados de una manera singular. Cada hexagrama impar va seguido de otro hexagrama que es su inverso o su complemento. Si el hexagrama impar tiene simetría bilateral (es el mismo por arriba que por abajo), su complementario es el hexagrama siguiente. Si no hay simetría bilateral, queda invertido.

¿Hay algún tipo de orden matemático que determine la forma en la que los pares de hexagramas se suceden uno a otro? Ese problema no se ha resuelto todavía. De vez en cuando, algún estudioso del *I Ching* anuncia el descubrimiento de un esquema matemático que se hallaría en la base de la colocación de los pares; ahora bien, si se analiza con rigor resulta que a tal fin se necesitan tantos presupuestos que, en realidad, el orden se ha englobado dentro de esos presupuestos, antes que se dedujera del análisis. Por lo que se sabe, los pares de la serie Rey Wen tienen un orden aleatorio, y no hay bases conocidas para determinar qué miembro del par debe preceder al otro.

Hasta el siglo XI, los eruditos chinos no descubrieron una manera muy simple y atractiva de ordenar los hexagramas. Esta colocación se atribuye a Fu Hsi *(véase la ilustración superior de la página 114).* La línea inferior en blanco representa el "t'ai chi", el estado del universo cuando estaba "sin forma y vacío" (como se lee en el *Génesis* 1, 2). Este caso indiferenciado se divide en dos mitades en la línea 1, la yin (en color) y la yang (en negro). En la línea 2 vemos el yin dividido en yin y yang, y análogamente el yang. Esta división binaria continúa hacia arriba por 6 veces.

Ahora, el esquema da automáticamente todos los poligramas de órdenes 1 hasta 6. Si se dividen las filas 1 y 2 verticalmente en cuatro partes iguales, y se substituyen en cada parte el color por líneas discretas (yin), se obtienen los cuatro diagramas. Las filas 1, 2 y 3, divididas verticalmente en 8 partes iguales, dan los ocho trigramas. Las filas 1, 2, 3, 4, en 16 partes, dan los 16 tetragramas; las filas 1, 2, 3, 4, 5, en 32 partes, dan los 32 pentagramas; por último las filas 1, 2, 3, 4, 5 y 6, en 64 partes, dan los 64

8 PI	7 SHIH	6 SUNG	5 HSÜ	4 MÊNG	3 CHUN	2 K'UN	1 CH'IEN
16 YÜ	15 CH'IEN	14 TA YU	13 T'UNG JÊN	12 P'I	11 T'AI	10 LÜ	9 HSIAO CH'U
24 FU	23 PO	22 PI	21 SHIH HO	20 KUAN	19 LIN	18 KU	17 SUI
32 HÊNG	31 HSIEN	30 LI	29 K'AN	28 TA KUO	27 I	26 TA CH'U	25 WU WANG
40 HSIEH	39 CHIEN	38 K'UEI	37 CHIA JÊN	36 MING I	35 CHIN	34 TA CHUANG	33 TUN
48 CHING	47 K'UN	46 SHÊNG	45 TS'UI	44 KOU	43 KUAI	42 I	41 SUN
56 LÜ	55 FÊNG	54 KUEI MEI	53 CHIEN	52 KÊN	51 CHÊN	50 TING	49 KO
64 WEI CHI	63 CHI CHI	62 HSIAO KUO	61 CHUNG FU	60 CHIEH	59 HUAN	58 TUI	57 SUN

Colocación Rey Wen de los 64 hexagramas del **I Ching**

hexagramas. La ilustración de la página 115 muestra los hexagramas en el tradicional orden Fu Hsi o "natural". De derecha a izquierda, y de abajo arriba, los hexagramas corresponden a los que da el esquema Fu Hsi leído de izquierda a derecha.

Ahora podemos entender por qué Leibniz, quien creía haber inventado el sistema binario a finales del siglo XVII, se sorprendió mucho al conocer por primera vez la sucesión Fu Hsi, a través del Padre Joachim Bouvet, un jesuita misionero en China. Sustituyendo por un 0 cada línea continua, y por un 1 cada línea discreta, y tomando los hexagramas por orden, leyéndolos de abajo arriba en cada uno, se obtiene la sucesión 000000,000001, 000010, 000011, ..., 111111; que no es otra que la de los números del 0 al 63, expresados en notación binaria.

Tanto Leibniz como el Padre Bouvet estaban convencidos de que Fu Hsi, bajo inspiración divina, había descubierto la aritmética binaria; pero no existe la mínima prueba de ello. En el siglo XI, los estudiosos del *I Ching* no habían hecho otra cosa que descubrir una manera natural de colocar los hexagramas. Hasta los tiempos de Leibniz no se reconoció la sucesión Fu Hsi como isomórfica a la notación aritmética binaria.

Puesto que las potencias del 2 se manifiestan continuamente en las estructuras matemáticas y en las físicas, no es de extrañar que los eruditos chinos pudieran aplicar los 64 hexagramas a casi todo, desde las estructuras de los cristales hasta el sistema solar y el cosmos. Z. D. Sung, en un divertido librito *The Symbols of Yi King* (Shanghai, The China Modern Education Company, 1914), cuenta cómo un día que estaba haciendo girar en la mano una caja de cerillas (para simular la rotación de la tierra alrededor del sol), de repente le vino a la mente un modo natural de generar los ocho trigramas en los vértices de un cubo.

Sean las tres coordenadas cartesianas de un cubo unidad, x, y, z, que indican el primero, segundo y tercer dígitos de un número binario trídigito. El vértice origen de coordenadas se designa con 000. Los otros vértices se designan con números binarios tridígitos desde el 0 al 7, donde 0 y 1 indican la distancia del vértice al origen, en cada una de las direcciones de las coordenadas. Los ocho números corresponden, obviamente, a los ocho trigramas, con los trigramas complementarios en los vértices diametralmente opuestos del cubo *(véase la ilustración de la página 116)*. Por un procedimiento similar, los vértices de hipercubos unidad generan poligramas de orden superior. Los 64 hexagramas corresponden a los números binarios de seis dígitos, situados en los vértices de un hipercubo de seis dimensiones.

En lugar de entrar en dimensiones superiores, Sung divide el cubo en 64 cubos más pequeños, a los cuales identifica con los 64 "modos" del silogismo clásico (la premisa mayor, la premisa menor y la conclusión de un silogismo pueden ser cada una de 4 formas diferentes, dando 64 modos posibles). Sung ignoraba probablemente que esta conversión ya la hizo C. Howard Hinton, en 1904, en su libro *The Fourth Dimension* páginas 90-106). Hinton se adentra de una forma curiosa en el hiperespacio; considerando las cuatro "figuras" de cada silogismo (una antigua división

TRIGRAMA	NOMBRE	IMAGENES	RASGOS	PARENTESCO	PARTES DEL CUERPO	ANIMALES
☰	CH'IEN	CIELO FRIO	FUERTE FIRME CLARO	PADRE	CABEZA	CABALLO
☷	K'UN	TIERRA CALOR	DEBIL FLEXIBLE OSCURO	MADRE	VIENTRE	ZORRO
☳	CHEN	TRUENO PRIMAVERA	ACTIVO MOVIL ESTIMULANTE	EL MAYOR DE LOS VARONES	PIE	DRAGON
☵	K'AN	AGUA LUNA INVIERNO	PELIGROSO DIFICIL ENVOLVENTE	EL SEGUNDO VARON	OREJA	CERDO
☶	KEN	MONTAÑA	RELAJANTE TENAZ INMOVIL	EL MENOR DE LOS VARONES	MANO	PERRO
☴	SUN	VIENTO BOSQUE	DISTINGUIDO PENETRANTE FLEXIBLE	LA MAYOR DE LAS HIJAS	MUSLO	PAJARO
☲	LI	FUEGO SOL RELAMPAGO VERANO	HERMOSO DEPENDIENTE AJUSTADO	LA HIJA SEGUNDA	OJO	FAISAN
☱	TUI	LAGO PANTANO LLUVIA OTOÑO	ALEGRE SATISFECHO COMPLACIENTE	LA HIJA MENOR	BOCA	OVEJA

Los 8 trigramas con algunos de sus significados

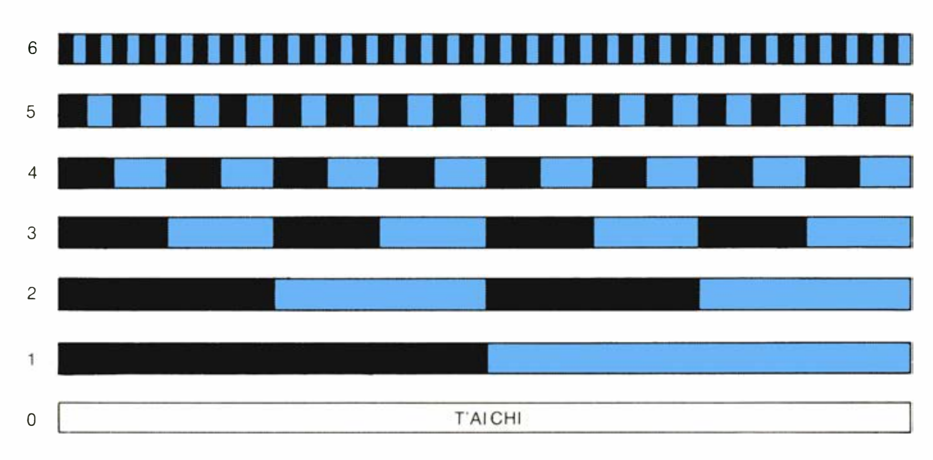
basada en la ordenación de sujeto, predicado y términos complementarios), obtiene 256 variedades, que identifica con las 256 celdas de un hipercubo de 4 por 4 por 4 por 4. Las celdas que corresponden a silogismos tradicionalmente válidos se pintan en negro y luego el hipercubo se proyecta sobre un cubo ordinario 4 por 4 por 4. Se observa entonces que las celdas negras estarían colocadas simétricamente alrededor de un vértice, excepción hecha de una celda que debería ser negra y no lo es. Esto llevó a Hinton al descubrimiento de que el silogismo anómalo puede ser válido, si se da una interpretación más liberal a los silogismos, interpretación en la cual se cuantifique, además del sujeto, el predicado.

Pero volvamos al *I Ching.* El libro (aparte de sus "Diez alas", que son apéndices incorporados por metafísicos confucianos) consta esencialmente de los 64 hexagramas, seguido cada uno de una breve explicación del símbolo y de seis "glosas". Cuando el libro sirve de oráculo, se ha de seleccionar al azar uno de los seis hexagramas, y ello de tal forma que las reglas digan cómo transformar el hexagrama escogido en un segundo hexagrama.

El procedimiento más antiguo de selección, seguido todavía por aquellos que consideran con extrema seriedad el uso adivinatorio del *I Ching,* exige disponer de 50 tallos de aquilea o mielenrama de 30 a 60 cm de longitud cada uno. Si no pueden conseguirse los tallos de mielenrama, servirán igualmente 50 palillos de madera. Deben guardarse en un recipiente cerrado, en un lugar cuya altura mínima sea los hombros de un hombre. El *I Ching,* cuidadosamente envuelto en seda limpia, se guarda junto a los palillos.

El libro no debe consultarse nunca a la ligera; si se le pregunta algo frívolo o con escepticismo, da una respuesta frívola o carente de sentido. Hay que estar completamente relajado, de cuerpo y de mente. Es esencial no pensar en nada durante todo el tiempo que dure la ceremonia, salvo en lo que se está preguntando.

Supongamos que se le va a preguntar algo al *I Ching,* y echando los palillos. El primer paso a dar es desenvolver la seda que cubre el libro, extenderla sobre la mesa y poner el libro encima (la tela protege al *I Ching* de las superficies impuras). Al lado del libro se coloca un braserillo de incienso y la caja que contiene los palillos. De espaldas al sur, arrodillados, se hacen tres inclinaciones profundas, tocando el suelo con la frente; luego, siguiendo arrodillados, se pasan los 50 palillos, por tres veces, a través del humo del incienso, sosteniéndolos horizontalmente y moviendo la mano en círculo en el sentido de las agujas del reloj. Se devuelve un palillo a la caja. Ese palillo no realizará ningún otro papel en la ceremonia.



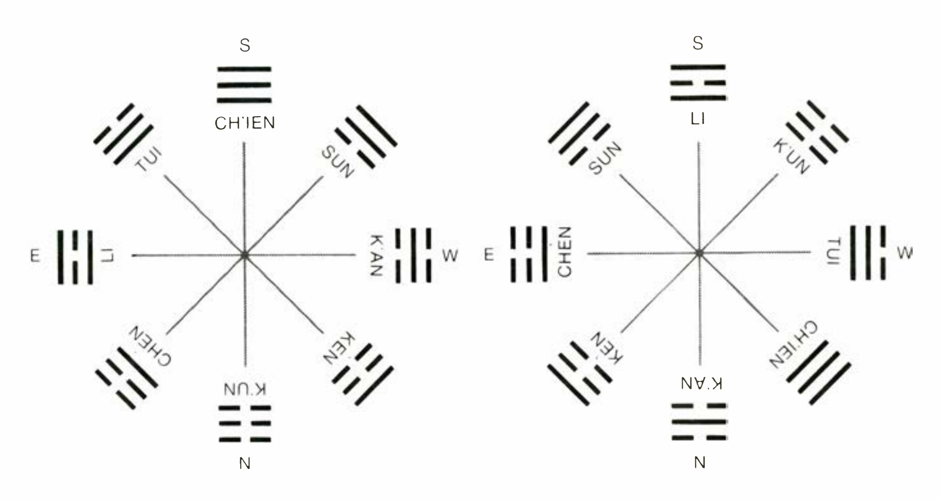
Forma en que seis divisiones yin-yang generan los 64 hexagramas

Se colocan los 49 palillos restantes sobre la tela y, luego, con la mano derecha se les divide rápidamente, al azar, en dos montones. Llamemos al montón de la izquierda montón *A*, y al otro montón, *B*. Se toma un palillo de *B* y se le introduce entre el anular y el meñique de la mano izquierda. Con la mano derecha se apartan del montón *A* palillos, de cuatro en cuatro, hasta que queden uno, dos, tres o cuatro palillos. Se disponen estos palillos entre el dedo medio y el anular de la mano izquierda. A continuación, se van restando del montón *B* cuatro palillos cada vez, hasta que queden uno, dos, tres o cuatro. Se colocan entre los dedos índice y medio de la mano izquierda (este último paso puede abreviarse; puesto que la suma de los dos restos debe ser 0, módulo 4, el segundo resto se calcula fácilmente a partir del primero). En la mano izquierda habrá ahora cinco o nueve palillos (las combinaciones posibles son 1,1,3; 1,2,2; 1,3,1; y 1,4,4). Todos estos palillos se colocan a un lado.

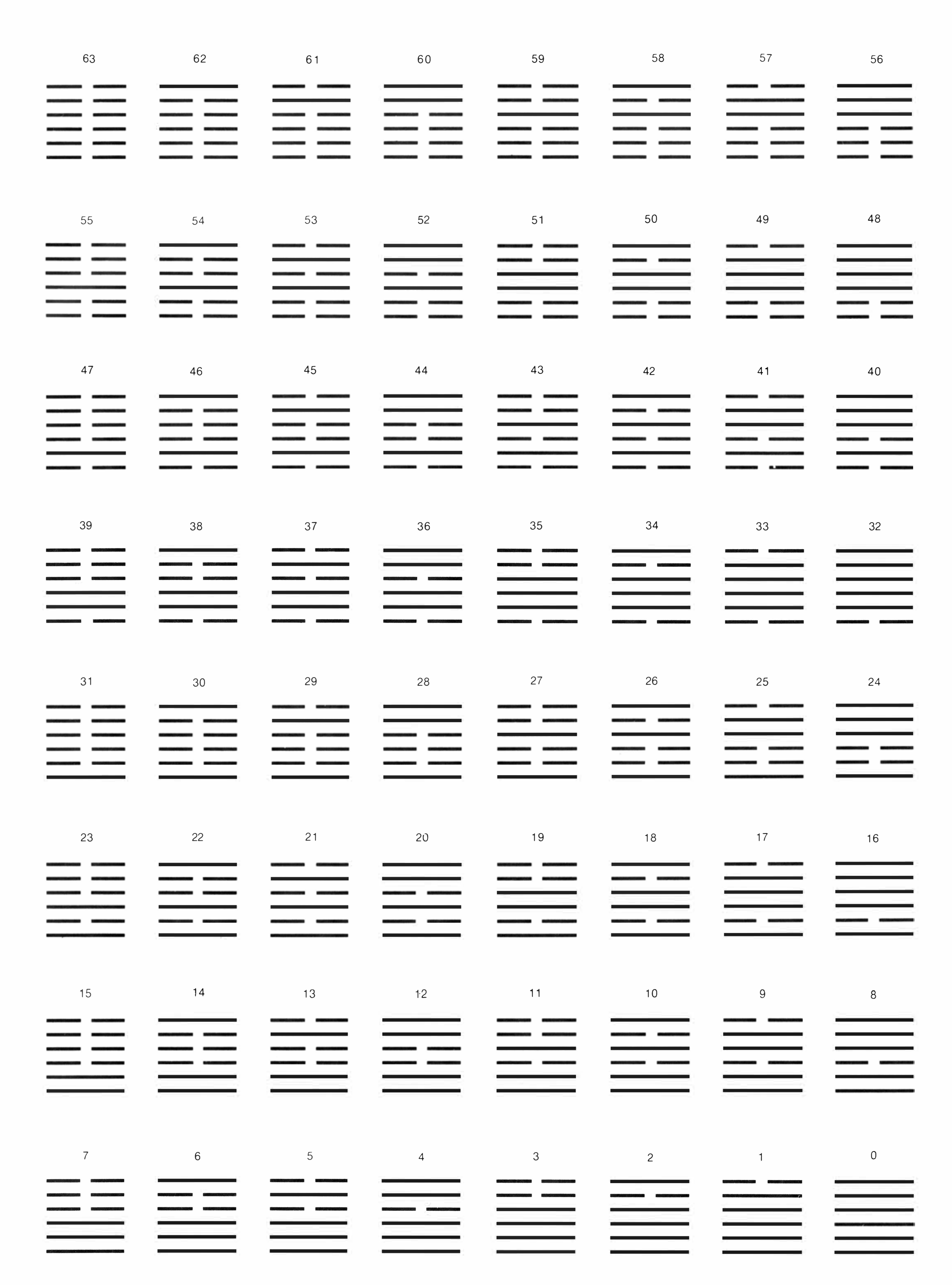
Se agrupan los palillos restantes y se repite exactamente con ellos el mismo procedimiento de división, empezando por la repartición al azar en dos montones. Al final, en la mano izquierda quedarán cuatro u ocho palillos (las combinaciones posibles son 1,1,2; 1,2,1; 1,3,4; y 1,4,3.) Se les pone a una lado, junto al grupo ya separado previamente.

Se agrupan los palillos restantes, y se repite el proceso de división por tercera vez. En la mano izquierda habrá, de nuevo, cuatro u ocho palillos. Se ponen a un lado, junto a los grupos ya colocados.

El número de palillos que ahora quedan será 24, 28, 32 ó 36. Hay que separarlos en grupos de cuatro (es decir, se divide por cuatro el número total restante). El cociente será 6, 7, 8 ó 9. Estos cuatro dígitos son los números rituales que indican el carácter de la línea inferior del hexagrama. Si el dígito es par (6 u 8), la línea será yin (discreta); si es impar (7 o 9), la línea será yang (continua). Pero los números rituales dicen aún más. Los números 7 y 8 significan que la línea (sea yin o yang) es una línea



La colocación de los trigramas atribuida a Fu Hsi (izquierda) *y la atribuida al Rey Wen* (derecha)



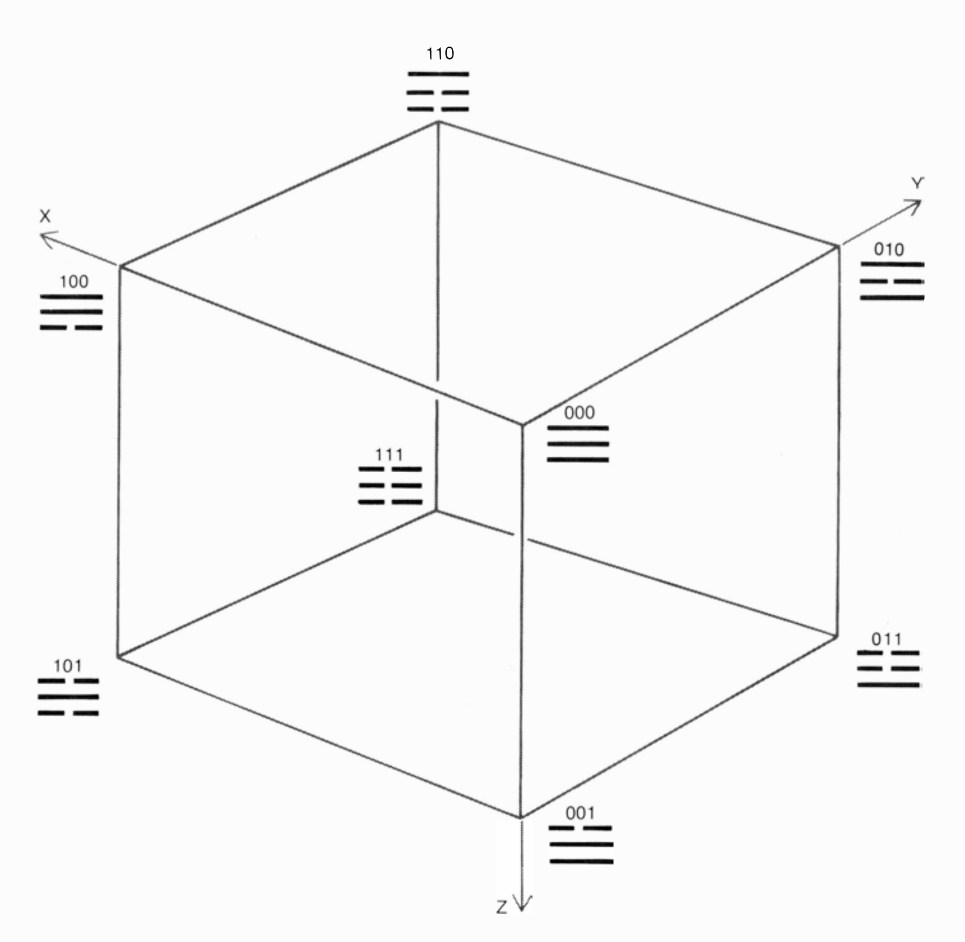
Secuencia de Fu Hsi que corresponde con los números binarios del 0 al 63

estable, que no puede alterarse. El 6 y el 9 indican una línea "móvil", que puede cambiarse (por razones que pronto se explicarán) por su opuesta.

Los 49 palillos se vuelven a agrupar y se repite todo el ritual para obtener la segunda línea del hexagrama, empezando por abajo. Cuatro repeticiones más, y tendremos las cuatro líneas restantes. Toda la ceremonia, realizada sin prisa, dura unos 20 minutos.

Se busca entonces el hexagrama escogido en el *I Ching,* y se estudia detenidamente el texto que lo acompaña. El texto responderá a la pregunta, aconsejando de acuerdo con la situación. Si las seis líneas del hexagrama son estables, se habrá llegado al final del proceso. Pero si una o más líneas son móviles, se cambian por sus opuestas y se considera el nuevo hexagrama. El comentario dirá qué es lo que se puede esperar en el futuro si se sigue el consejo del primer hexagrama.

Tras escribir el hexagrama o los dos hexagramas, y haber leído y meditado sobre los pasajes correspondientes del *I Ching,* se enciende otra barrita de incienso, se hacen otras tres inclinaciones profundas, se devuelven los palillos a su caja, se envuelve de nuevo el *I Ching* en su seda, y, por fin, se deposita el libro y los palillos en su lugar habitual.



Trigramas generales por un cubo

Quienes se consideren demasiado perezosos para seguir el antiguo ritual de los palillos pueden seguir un método sencillo de actuación, que se hizo popular en China durante siglos. Se necesitan seis monedas idénticas, preferentemente monedas chinas antiguas con agujeros cuadrados. Deben guardarse limpias, sin tocarlas de su caja, excepto cuando se haya de consultar el *I Ching.* Al principio, se procede según el mismo ritual seguido con los palillos: se saluda humildemente, se hincan las rodillas, se pasan las monedas por el incienso, etc... Se agitan las monedas con las manos juntas y se sueltan a la vez sobre la tela. Se decide previamente qué cara de las monedas sea yin y qué cara yang y se consulta la siguiente tabla para determinar si la tirada da 6, 7, 8 o 9.

Tres yin =6 (una línea yin móvil).
Dos yin, un yang =7 (línea yin estable).

Dos yang, un yin =8 (una línea yang estable). Tres yang =9 (una línea yang móvil) (si se asigna el valor 2 a la cara yin y el valor 3 a la cara yang, la suma de los tres valores será el número ritual deseado).

Calculadas las probabilidades del procedimiento de los palillos y las del procedimiento de las monedas aparece una sutil diferencia entre ambos métodos de adivinación. En lo que se refiere a la elección del primer hexagrama, los métodos son virtualmente idénticos, pero en la elección del segundo hexagrama las probabilidades varían. No es difícil demostrar que, en ambos procedimientos, la probabilidad de obtener una línea discreta en cada uno de los seis pasos es 1/2; la misma que la de obtener una línea continua. (Se supone que cada vez que los palillos se dividen al azar en dos montones *A* y *B*, y *A* se reduce a uno, dos, tres o cuatro palillos, las probabilidades para cada uno de los cuatro resultados son iguales. Lo cual no es estrictamente cierto, pero se apartan tan ligeramente de la igualdad que afecta muy poco a los resultados finales.) Cada hexagrama tiene por tanto la misma probabilidad de ser el elegido. Los dos procedimientos son semejantes: otorgan una probabilidad de 1/4 respecto a que la línea sea móvil. Puesto que hay seis líneas, 6/4, es decir, 1 $^1/_2$ líneas del hexagrama, por término medio, serán móviles.

Al usar monedas, la probabilidad de que una línea discreta cambie es la misma (1/4) que la probabilidad de que cambie una línea continua; del mismo modo, la probabilidad de que cada tipo de línea se mantenga estable será 3/4. Pero no es este el caso cuando se utilizan palillos. La probabilidad de que una línea discreta cambie es 1/16 frente a los 3/16 de una línea continua (o bien se tendrán respectivamente las probabilidades de 7/16 frente a 5/16 de que tales líneas permanezcan estables). En otras palabras, al lanzar los palillos, la probabilidad de que cambie una línea continua es tres veces mayor que la que lo haga una línea discreta. Un hexagrama tiene la misma probabilidad que cualquier otro de ser escogido el primero, pero cuanto mayor sea el número de líneas discretas que tenga, tanto mayor es la probabilidad de que aparezca como segundo hexagrama. Los puristas que se han opuesto al método de las monedas disponen de sólidos argumentos matemáticos. El ritual de los palillos no sólo no da pie a consultas frívolas sino que su asimetría produce un conjunto de probabilidades más interesante. No diremos nada de corrupciones tan desvergonzadas como la obtención de números rituales a partir de billetes de banco, matrículas de coches, números de teléfonos, etcétera.

Para los lectores que deseen experimentar con el *I Ching,* mi primera recomendación es que lean la traducción alemana de Richard Wilhelm *I Ching, das Buch der Wandlungen,* Jena 1924; reimp. Düsseldorf, 1956) traducido al inglés

por C. F. Baynes, editado actualmente en edición de lujo por la Princeton University Press; existen otras dos magníficas traducciones en rústica, de James Legge (Dover) y de John Blofeld.

El volumen de Wilhelm-Baynes incluye el famoso prólogo de Jung, en que explica el poder oracular del *I Ching* mediante su teoría de la "sincronicidad", teoría defendida por Arthur Koestler en su reciente libro *The Roots of coincidence*. Según Jung, entre las predicciones del *I Ching* y los importantes acontecimientos de la vida real no media una razón causal en el sentido occidental de la ciencia. Predicciones y hechos se hallan relacionados de una manera "causal" en el sentido de la metafísica oriental, como partes de un vasto designio cósmico que, a pesar de estar más allá del alcance de la ciencia, es, no obstante, parcialmente accesible para el inconsciente de la persona que echa los palillos. Los 64 hexagramas, y sus significados, son arquetipos de Jung, profundamente grabados en el inconsciente colectivo de la humanidad. En castellano, disponemos de la edición de MirKo Laner (Barral Editores, Barcelona, 1975[7]).

Los escépticos obstinados que ponen a prueba el *I Ching* alegan siempre que el texto es tan ambiguo que, con independencia de los hexagramas seleccionados, siempre se puede dar una interpretación de los mismos de manera que parezca tener algo que ver con la pregunta. Realmente, el campo para la interpretación intuitiva es tan amplio que, en China, antes de Mao (no sé cómo será hoy) había muchos intérpretes profesionales del *I Ching,* de cuyos servicios podía disponerse, previo pago, en las esquinas de las calles, en las ferias y en las plazas de los mercados. A buen seguro, una razón de la popularidad del método de las monedas fue que maximizaba las ganancias de los adivinos al concluir antes sus lecturas.

¿Y si las predicciones del *I Ching* no se cumplen? Pues, en este caso, quizás el texto no se interpretó correctamente, o tal vez no se estuviera con el estado de ánimo adecuado al tirar los palillos o las monedas. Además, el futuro no se halla determinado de un modo absoluto, y el *I Ching*, como las estrellas en la astrología, sólo indica tendencias probables.

Se avisa a los devotos del ocultismo que no hayan consultado todavía el *I Ching* y anhelen poderosos y misteriosos medios mágicos que los consejos de este viejo libro pueden producir un impacto psicológico mucho mayor que los de los simples astrólogos, quirománticos y adivinos de la bola de cristal o posos de té.

Taller y laboratorio

Un sencillo experimento para iniciarse en una poderosa técnica analítica: la cromatografía

Durante los últimos años, millares de aficionados han descubierto el interés de la cromatrografía, o "análisis por adsorción", como se le ha denominado en algunas ocasiones. A pesar de que el botánico ruso Michael Tswett describió el método cromatográfico en 1906, no se generalizó su uso, ni siquiera entre los profesionales, hasta 1930.

Ninguna descripción del método cromatográfico ha sobrepasado en claridad y concisión a la dada originalmente por Tswett: "Si una solución de clorofila en éter de petróleo se filtra a través de una columna de un adsorbente (yo uso generalmente carbonato cálcico, que comprimo fuertemente en el interior de un estrecho tubo de cristal), los pigmentos se separan de arriba abajo, de acuerdo con la secuencia de adsorción, en varias zonas coloreadas, como los rayos de luz en el espectro, de forma que los diferentes componentes de una mezcla de pigmentos se resuelven en una columna de carbonato cálcico de acuerdo con una ley y pueden ser valorados sobre ella cualitativa y cuantitativamente. A tal preparación yo le llamo un cromatograma, y, al método correspondiente, el método cromatográfico. Es evidente que el fenómeno de adsorción descrito no se restringe a los pigmentos de la clorofila, y se ha de suponer que todas las clases de compuestos químicos, coloreados y no coloreados, están sujetas a las mismas leyes."

En esencia, la cromatografía sólo requiere tres aparatos: un recipiente para contener la muestra, la columna cromatográfica y un segundo recipiente para recoger el líquido usado a medida que gotea por el final del tubo. Después de rellenar la columna con un material adsorbente, se coloca en la parte superior del tubo una porción de la solución de muestra y se deja que descienda a través de la columna, entre una décima y una cuarta parte de su longitud. Al hacer esto, normalmente se forma una banda uniforme del color característico de la solución que se está estudiando. Entonces se hace pasar disolvente de lavado a través de la columna, y comienza el proceso de separación. Cada substancia tiene unas afinidades características por el disolvente y por el adsorbente. Los expertos suelen denominar a esta propiedad la "actividad" del adsorbente o del disolvente. La relación de actividades determina la posición que una substancia particular ocupará en la columna respecto a las restantes de la mezcla de la que está siendo separada. Las substancias menos solubles y con mayor afinidad por el adsorbente quedan adheridas a las partículas superiores de la columna. Aquellas de menor afinidad por el adsorbente serán lavadas más rápidamente, dependiendo su desplazamiento de la adsorción relativa de cada substancia. Las substancias así separadas formarán en la columna un patrón característico de bandas. Por ejemplo, los extractos preparados a partir de algunas hojas verdes presentan más de veinte bandas de colores diferentes, desde verde oscuro hasta blanco, pasando por varios tonos de anaranjado, rosa, amarillo y violeta suave. Estos colores corresponden a las diversas xantofilas, flavoxantinas, luteínas, carotenos y pigmentos afines.

La operación de lavar la columna con disolvente limpio se conoce como "desarrollo" de la columna. Cuando un solvente nuevo circula a través de una columna, algunas moléculas se desprenden del adsorbente, se incorporan a la solución y bajan a regiones de concentración menor. Allí son readsorbidas. Las actividades del solvente y del adsorbente varían según la concentración de la substancia que se está analizando; así, al repetirse el proceso expuesto, una substancia puede entrar y salir de la solución varias veces en el curso de su descenso a través de la columna. Al principio, las bandas son estrechas y están agrupadas cerca del principio del tubo. A medida que avanza el desarrollo de la cromatografía, todas las bandas descienden hacia el final del tubo, separándose más claramente y haciéndose más anchas. Un cromatograma completamente desarrollado presenta una serie de bandas definidas y bien separadas, cuya diferente anchura es proporcional a la cantidad que de cada substancia había en la mezcla inicial.

Las substancias ya separadas y purificadas pueden ser extraídas por uno de estos dos procedimientos: o bien prosiguiendo el lavado con disolvente, hasta que las sucesivas bandas salgan por el final de la columna, o bien empujando el cilindro del adsorbente para sacarlo del tubo, cortando seguidamente las bandas con una cuchilla y extrayendo la substancia de cada trozo con un disolvente. Si se succiona el cromatograma para secarlo, es más fácil sacarlo del tubo. Algunos autores extraen el adsorbente del tubo separando las bandas una a una con una fina espátula.

Se han probado miles de adsorbentes y solventes distintos. La selección de la combinación más efectiva para cada propósito suele ser el resultado de repetidas pruebas. La siguiente lista de adsorbentes y solventes, con los cuales se pueden separar la mayoría de las mezclas que pueda preparar el aficionado, fué descrita por William T. Beaver, un estudiante de la Universidad de Princeton, y mereció un galardón de la Westinghouse Science Talent Search. Los adsorbentes se presentan por orden aproximado de actividad de mayor a menor; los solventes en el orden inverso:

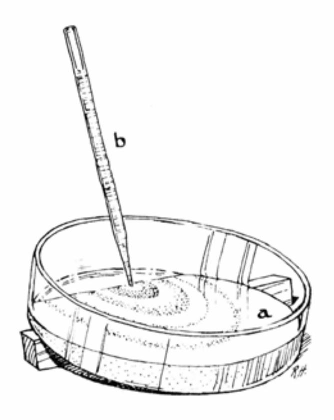
ADSORBENTES	SOLVENTES
1 – Alúmina activa	1 – Eter de petróleo
2 – Carbón vegetal	2 – Tetracloruro de carbono
3 – Magnesio	3 – Disulfuro de carbono
4 – Gel de sílice	4 – Eter
5 – Cal	5 – Acetona
6 – Carbonato magnésico	6 – Benceno
7 – Carbonato cálcico	7 – Alcohol metílico o etílico
8 – Carbonato sódico	8 – Agua
9 – Talco	9 – Acidos orgánicos
10 – Azúcar en polvo	10 – Soluciones acuosas de ácidos o bases.

A veces, en lugar de un solo solvente, es preferible utilizar varios, ya sea en combinación o uno tras otro. Por ejemplo, puede mezclarse una pequeña cantidad de benceno con un solvente de baja actividad, como el éter de petróleo, para acelerar el desarrollo de las bandas. Sin embargo, es necesario comprobar que el solvente seleccionado no sea tan activo

que lave inmediatamente las bandas de la columna. Una vez se han cortado en bloques las bandas del cilindro de adsorbente, se los puede tratar con un solvente de gran actividad para conseguir una extracción rápida y completa de las substancias principales. A este proceso se le denomina "elución" y al solvente o combinación de solventes utilizados "eluyentes". La mayoría de los adsorbentes y solventes más corrientes son baratos. Algunos se encuentran al alcance de todos, incluso en la propia casa. Beaver aconseja al principiante que adquiera los productos para cromatografía en un establecimiento de productos químicos. Los productos caseros estarán probablemente contaminados, y una pequeña cantidad de una materia extraña puede confundir los resultados. La cromatografía es una técnica extremadamente sensible.

Como dice Beaver, "el hecho de que existan pocas reglas establecidas sobre cromatografía es lo que la hace recomendable como afición. Ni siquiera el profesional más avanzado puede aconsejar un procedimiento preciso y rápido para fabricar y manejar una columna cromatográfica. El campo es tan nuevo que está abierto a todos los interesados. El aficionado tiene la gran oportunidad de hacer aportaciones importantes a la técnica".

El diámetro interno de la columna de vidrio puede medir desde unos milímetros hasta varios centímetros, dependiendo de la textura del adsorbente, la naturaleza de las substancias que van a ser adsorbidas, la cantidad de material disponible y otras consideraciones semejantes. La mayoría de investigadores prefiere usar columnas de uno o dos centímetros de diámetro interior. Normalmente las columnas son unas diez veces más largas que anchas. Para la separación de algunos isótopos, sin embargo, se han usado tubos delgados de varias decenas de metros de longitud. El final del tubo se tapa con una bola de algodón o de lana para proporcionar un soporte al adsorbente. Tubos de características similares pueden obtenerse en la mayoría de establecimientos de productos químicos, pero también pueden fabricarse fácilmente en casa con tubo de vidrio.

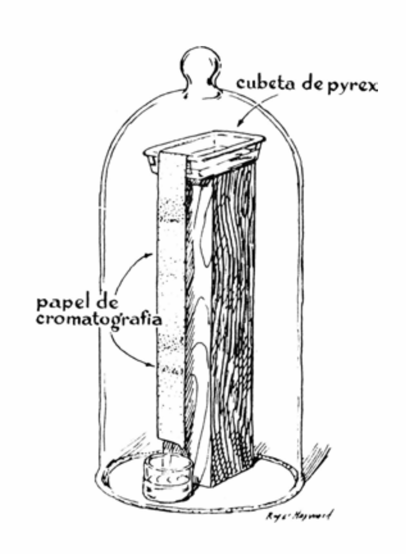


Método conveniente para escoger los solventes y adsorbentes.

La mayor dificultad práctica para el principiante es la de rellenar uniformemente la columna. Si el adsorbente no está uniformemente distribuido, lo más probable es que las bandas sean poco definidas y se solapen. Tswett iba introduciendo poco a poco el adsorbente en forma de polvo seco, presionándolo ligeramente, hasta lograr la longitud deseada. Experiencias posteriores han aconsejado modificar este procedimiento en varias formas. Después de haber colocado y apretado una de las capas de polvo seco en la columna, se puede disgregar ligeramente su superficie con una espátula, con lo que la siguiente se acoplará mejor, produciendo un resultado más uniforme. Un palo de madera normal y corriente, con un extremo plano y algo más estrecho que el diámetro interior del tubo, constituye un buen instrumento para taponar. Algunos adsorbentes se asientan satisfactoriamente en la columna si se hace vibrar ligeramente el tubo mientras se le llena lentamente.

La mayoría de los investigadores seleccionan los solventes y adsorbentes más apropiados para cada trabajo de acuerdo con técnicas conocidas y estandardizadas. Uno de los procedimientos más populares consiste en colocar una cucharadita del adsorbente en una cápsula de petri, inclinándola para que el contenido forme una especie de cuña sobre el fondo. La mezcla que se va a probar se disuelve en un disolvente débil y se colocan unas cuantas gotas de éste sobre el borde fino del adsorbente con una micropipeta. En estas condiciones se aplican varios solventes y combinaciones de ellos, comprobándose su mayor o menor actividad *(véase la ilustración de la primera columna)*.

Los aficionados que quieran adentrarse en este campo recurrirán sin duda antes o después a la cromatografía sobre papel, técnica que facilita mucho las cosas. La "columna" es en este caso una tira u hoja de papel, que se introduce en el seno de una atmósfera saturada para evitar la evaporación. Tras humedecer el papel con disolvente, se aplica una gota de la solución que va a ser analizada en el borde superior de la hoja, o en una de sus esquinas. Seguidamente se dobla y



Cromatografía sobre papel (realizada en una cámara para evitar la evaporación de la muestra).

se introduce en una cubeta que contenga el disolvente que se va a utilizar para su desarrollo. El solvente desciende por capilaridad por la hoja así dispuesta, transportando las substancias a resolver, que son adsorbidas formando manchas a lo largo del papel. Estas manchas corresponden a las bandas de la columna convencional. Cuando el desarrollo hace que la mancha inferior se halle cerca del final de la hoja, puede sacarse ésta del solvente, girarla 90 grados e introducirla de nuevo en él. Cada mancha se convierte entonces en el punto de origen de un nuevo cromatograma. Si las fracciones desarrolladas están formadas por la mezcla de dos o más substancias, aparecerán ahora separadas a lo largo de la hoja. Lo que se obtiene entonces es una "cromatografía en dos dimensiones" *(véase la ilustración de esta tercera columna)*.

Tswett comparaba las bandas que aparecían en su columna con los rayos de luz coloreada que emergen de un prisma formando un arco iris. La cromatografía en dos dimensiones refuerza la analogía al volver a someter cada mancha a un segundo análisis, incrementando así la resolución del método, de forma parecida a como ocurre en óptica, cuando los físicos hacen pasar un rayo coloreado que proviene de un prisma, a través de un segundo prisma. Muchos aficionados utilizan la técnica sobre papel como un método para determinar cualitativamente las fracciones de una mezcla, y luego usan la columna convencional para la determinación cuantitativa.

Tal como predijo Tswett, la técnica cromatográfica permite separar tanto las substancias coloreadas como las no coloreadas. Ultimamente se está trabajan-

do mucho en cromatografía de substancias incoloras. Se han desarrollado gran número de técnicas para visualizarlas. Por ejemplo, la presencia de aminoácidos se detecta pulverizando ninhidrina sobre el extruido de adsorbente o sobre el papel de cromatografía, con lo que estas substancias, normalmente incoloras, adquieren un color rojo suave. Otras substancias emiten fluorescencia a la luz ultravioleta. Si se coloca una gota de tinta azul-negra corriente en el extremo de un papel de cromatografía y se desarrolla con alcohol, se forman a lo largo del papel una serie de bandas azuladas que representan los distintos compuestos de hierro y los diferentes colorantes de la tinta. Pero, bajo una lámpara de luz ultravioleta, el papel seco muestra muchas otras bandas, de colores rojos, anaranjados y verdes. Por medio de un segundo cromatograma, en el que se utiliza como control una substancia conocida, puede identificarse otra desconocida comparando la posición de las respectivas bandas sobre los cromatogramas. También pueden detectarse fracciones incoloras sobre una columna cromatográfica por la acción diferencial que estas substancias ejercen sobre la luz transmitida a través de la columna o por su polarización. Recientemente se han marcado substancias con isótopos radiactivos, que luego se detectan por procesos fotográficos, pero estas técnicas suelen estar lejos del alcance del aficionado medio.

La extracción de las clorofilas, carotenos y xantofilas de hojas de espinaca puede muy bien servir como introducción a la técnica cromatográfica. El primer experimento que vamos a describir, sugerido por Beaver, deberá realizarse en una habitación bien ventilada, ya que los disolventes son muy volátiles e inflamables. Además, los vapores de uno de ellos, el alcohol de madera, son venenosos.

En columnas fabricadas de tubo de vidrio de unos 10 milímetros de diámetro interior y unos 30 centímetros de longitud, con uno de sus extremos pulido al fuego y el otro ligeramente acampanado, para facilitar el llenado, se empaqueta el adsorbente (alúmina de la casa Merck, estandarelizada de acuerdo con Brockmann, de 80 a 200 de poro). El empaquetamiento se realiza poco a poco y agitando simultáneamente el tubo. La succión reduce el tiempo de desarrollo. El dibujo de Roger Hayward *(véase la ilustración de la izquierda)*, muestra cómo se ha de disponer la columna para utilizarla con un frasco de vacío. En la figura siguiente se presenta otra alternativa de alimentación sin necesidad de vacío.

Se maceran diez gramos de hojas de espinacas secas en 100 mililitros de alcohol de madera durante 24 horas. Seguidamente se filtra el material y el residuo se lava con 50 mililitros más de alcohol de madera. Este extracto se agita con 50 mililitros de éter de petróleo; se añaden 100 mililitros de agua y la solución resultante se coloca en un embudo de decantación. Cuando se han separado las fases, se descarta la inferior (de alcohol-agua) y se filtra la superior (éter de petróleo), que es la que contiene el extracto.

Se coloca la mitad aproximadamente de este extracto en la columna de alúmina y se desarrolla con benceno. La primera fracción que atraviesa la columna es una banda muy extrecha amarillo-naranja de caroteno. A ésta le siguen algunas bandas más anchas rosa-amarillentas, que son de xantofilas. Estos pigmentos le son familiares a todo el mundo, ya que dan el color de otoño a los bosques una vez que el frío ha destruido la clorofila. Pueden recogerse las distintas fracciones de estos pigmentos a medida que salen por el final de la columna y evaporarlas para concentrarlos. Los dos grupos, carotenos y xantofilas, pueden resolverse en sus componentes volviéndolos a disolver en unos mililitros de éter de petróleo, haciéndolos pasar por columnas nuevas y desarrollándolos con éter de petróleo-benceno o, si se necesita mayor actividad del eluyente, con benceno puro.

Las clorofilas forman una banda verde oscura, que también puede extraerse de la columna. Los pigmentos se lavan en 5 mililitros de alcohol de madera y se filtra la solución. Este filtrado se coloca en un embudo de decantación con 5 mililitros de éter de petróleo y se añaden 5 mililitros de agua. El éter de petróleo extrae las clorofilas y el agua forma con el alcohol otra fase que puede ser decantada. El extracto de éter de petróleo se lava varias veces con agua por este mismo procedimiento y se coloca luego en una columna empaquetada con azúcar en polvo (sacarosa) disuelta en éter de petróleo hasta formar una papilla. La columna asi preparada se desarrolla con éter de petróleo. Las clorofilas se separan en dos componentes: una banda amarillo-verde oscura de beta-clorofila,

Columna de cromatografía alimentada por gravedad.

Columna de cromatografía alimentada mediante el vacío.



que se sitúa en la parte superior de la columna, y otra verde-azulada de alfa-clorofila, colocada mucho más abajo.

Debido al decisivo papel que juega en la fotosíntesis, la gente piensa en la clorofila como el maravilloso pigmento de las plantas. Pero muchos aficionados a la cromatografía encuentran en los carotenos un interés igual o mayor. A diferencia de las clorofilas, que actúan como catalizadores, los carotenos desempeñan un papel químico directo tanto en los animales como en las plantas. Al parecer son esenciales para la fabricación de la molécula de la vitamina A, y tienen un papel importante en los mecanismos de la visión y en los sexuales. Como su nombre indica, pueden extraerse de la zanahoria *(Daucus carota)*.

Para extraer carotenos, se trituran unos 5 gramos de raíces de zanahoria secas en un mortero. Se añaden 50 mililitros de una mezcla de alcohol de madera y éter de petróleo a partes iguales. Se agita fuertemente la mezcla, se añaden 5 mililitros de agua y se coloca en un embudo de decantación. Los carotenos más los ésteres de xantofilas se concentran en la fase de éter de petróleo que se forma en la parte superior del embudo. Se separa esta fase y se concentra evaporando algo de líquido, dejando unos 20 mililitros. Seguidamente se añaden 3 mililitros de hidróxido sódico disuelto al 5 por ciento en alcohol de madera. Así se saponifican los ésteres de las xantofilas, con lo que pueden extraerse. Para ello se lava la mezcla varias veces con alcohol de madera disuelto en agua al 85 por ciento. Para eliminar los restos del alcohol, se lava luego varias veces con agua. Se separa el éter de petróleo del agua y se filtra. La solución amarillo-naranja que se obtiene está formada por el complejo de carotenos de las raíces iniciales disuelto en éter. Para separarlos, se coloca la mitad aproximadamente de la solución obtenida en una columna de alúmina y se desarrolla ésta con una mezcla de benceno y éter de petróleo en la relación de 1 a 3. Se obtendrán tres bandas bien definidas, que contienen, de arriba abajo, gamma-caroteno, beta-caroteno y alfa-caroteno. Se puede recuperar cada uno de los pigmentos lavándolos sucesivamente de la columna o extrayéndolos con alcohol de madera de las bandas separadas del adsorbente.

La cromatografía es un método de separación mucho más sutil que las técnicas tradicionales de destilación, precipitación con reactivos y cristalización. Es también un método de una simplicidad encantadora.

Libros

Ensayos de historia en la mecánica y los últimos adelantos en bioenergética

ENSAYOS DE HISTORIA DE LA MECÁNICA, de C. Truesdell, trad. de Juan C. Navascués y Enrique Tierno, Editorial Tecnos, Madrid 1975. El lector hispanohablante dispone de la traducción de la que se considera una de las mejores historias generales de la ciencia, la de Taton y cols. (París, 1957-1964; segunda edición en prensa). Pero halla un gran vacío en punto a la historia de las distintas ciencias empíricas trenzada con los textos de los principales protagonistas, si exceptuamos el caso de la medicina cuya historia nos ha sido contada con rigor de auténtica escuela.

Por eso hay que dar la bienvenida a la versión española de la obra compilada de Truesdell, en donde se recoge su interpretación de los escritos y de las líneas de investigación de la mecánica clásica. Del valor del contacto directo con las fuentes, él nos cita una muestra: "Conozco hombres jóvenes que han leído los escritos de Gibbs y Kelvin y Cauchy, incluso de Euler y Newton, no con el propósito de adornar un trabajo propio con una referencia temprana, ni de escribir una historia, sino en búsqueda de entendimiento y método, revelados ambos en las palabras directas de gigantes, no en las traducciones hechas por enanos".

La mecánica clásica se desgajó de la mecánica celeste por obra de Galileo, y de ella la mecánica cuántica por los descubrimientos de comienzos de siglo de Max Planck, Albert Einstein y Niels Bohr. Esta separación no debe entenderse siempre de forma absoluta. En particular, la mecánica celeste y la mecánica clásica anduvieron muy de la mano para explicar ciertos problemas cosmológicos (véase Callahan: "The curvature of Space in a Finite Universe" en *Scientific American,* agosto, 1976). La mecánica clásica engloba dos ramas más o menos fecundas: la estática o estudio de las fuerzas y del rozamiento, y la dinámica o estudio de la cinética, cinemática, momento y energía.

Nadie puede franquear el portal de la mecánica, de la física en general, sin un gran bagaje matemático. Ya pedía Platón que "nadie entrara allí que no fuera geómetra" cuando pretendía desentrañar, bajo la influencia pitagórica, los misterios de la armonía del universo. Por eso puede extrañar que comience este volumen con un extenso capítulo dedicado a Leonardo, que no andaba muy fuerte en matemáticas y a quien los problemas más elementales se le hacían ovillos enmarañados. Pero, en su defensa, él mismo alegaba: "sé construir puentes extremadamente ligeros y resistentes [...] sé de métodos para quemar y destruir los del enemigo [...] sé cómo sacar agua de los fosos cuando un lugar se encuentra sitiado [...] en caso de necesidad construiré grandes cañones [...]".

El ordenamiento sistemático de la mecánica comienza cuando integra el razonamiento matemático, en Newton, cuya obra máxima lo anunciaba: *Philosophiae Naturalis Principia Mathematica* (Principios matemáticos de la filosofía natural). La filosofía natural abarcaba entonces ciencias tan dispares como la física, la astronomía, la psicología, la botánica, etc.; faltaría más de un siglo para que se procediera a la separación de las distintas ciencias empíricas del tronco único de la filosofía. El capítulo siguiente está dedicado a las reacciones de la mecánica del barroco tardío a los éxitos, conjeturas, errores y fracasos contenidos en el *Principia* de Newton. Cauchy, matemático francés de primera fila, dará un nuevo impulso a la mecánica clásica con el *principio de esfuerzos,* cuyas piezas gestoras Truesdell descubre en la centuria anterior. Otros dos puntos interesantes de la mecánica, cuales son la ley del momento de la cantidad de movimiento y la teoría cinética de los gases, constituyen, desde una perspectiva histórica, sendos capítulos consecutivos a los anteriores. La obra termina con un alegato en pro de la eficacia del estudio de la mecánica racional, "rama independiente de la matemática cuyo objetivo es entender a la mecánica".

ENERGY TRANSFORMATION IN BIOLOGICAL SYSTEMS. CIBA FOUNDATION SYMPOSIUM 31. IN TRIBUTE TO FRITZ LIPMANN ON HIS 75 TH BIRTHDAY. Elsevier/Excerpta Medica/North Holland. La especificidad genética de la vida en evolución resulta un sistema maravilloso de almacenamiento y transmisión de información, un codicilo de la segunda ley de la termodinámica. La actividad de las formas vivas, entendidas ya no como especies sino como individuos, se gobierna por la primera ley de la termodinámica que rige el intercambio de energía. Primer principio que no es ni menos molecular ni menos importante. Detrás de la información subyacen las moléculas en disposición helicoidal de los ácidos nucleicos; detrás del intercambio de energía yace otra molécula, relativamente pequeña, de adenosintrifosfato (ATP). La energía del ATP se libera cuando la enzima adenosintrifosfatasa (ATPasa) activa la molécula.

A Lipmann se debe, desde los años inmediatamente anteriores a la Segunda Guerra mundial, la consideración de la importancia clave del "enlace fosfato de alta energía", al que simbolizó ($\sim$P). En su comunicación pasa revista a las "raíces de la bioenergética", ciencia que la han construido justamente quienes asistían a esa destacada reunión.

Tras la comunicación introductoria encontramos dos estudios en torno a la función y al origen de dos organelas celulares muy extendidas: los cloroplastos y las mitocondrias. En los cloroplastos, la energía lumínica se convierte en ATP y coenzima, y queda almacenada en enlaces carbono-carbono. Estos, a su vez, entran en combustión en las mitocondrias en presencia de oxígeno, y forman ATP de nuevo que permitirá el metabolismo celular. El flujo de hidratos de carbono y gases puede ser macroscópico; toda la atmósfera participa de este proceso. El circuito elaborado está muy lejos de conocerse al detalle; en particular, en lo que respecta a las estructuras de las membranas envolventes, que se disponen en razón de notables gradientes químicos. Se admite ampliamente que ambos tipos de organelas proceden de células procariotas libres, si bien el origen de los cloroplastos verdes a partir de algas azules y verdes se acepta con mucho mayor unanimidad. ¿Por qué estas organelas, al igual que todas las células procariotas, tardan tanto en evolucionar, se mantienen pequeñas y unicelulares? ¿Son inventores bioquímicos más que sistemas genéticamente elaborados? Lipmann aduce la respuesta de François Jacob: los procariotas no mueren. Y el ciclo de vida y muerte es esencial para la evolución.

Otro participante, F. R. Whatley, de Oxford, habla de un tipo de ameba que poco a poco se va autoenvenenando al

secretar ácido láctico. Carece de mitocondrias y, por lo tanto, es, de hecho, un organismo anaerobio. Enclaustrada en sí misma en pequeñas cavidades, alberga bacterias aeróbicas; el metabolismo de estos residentes aeróbicos le permite limpiar el ácido láctico perjudicial, aun cuando ellos fabriquen ATP para sí mismos, no para el hospedante. Hay una suerte de simbiosis anterior al origen postulado de las mitocondrias ("una historia evolutiva atractiva").

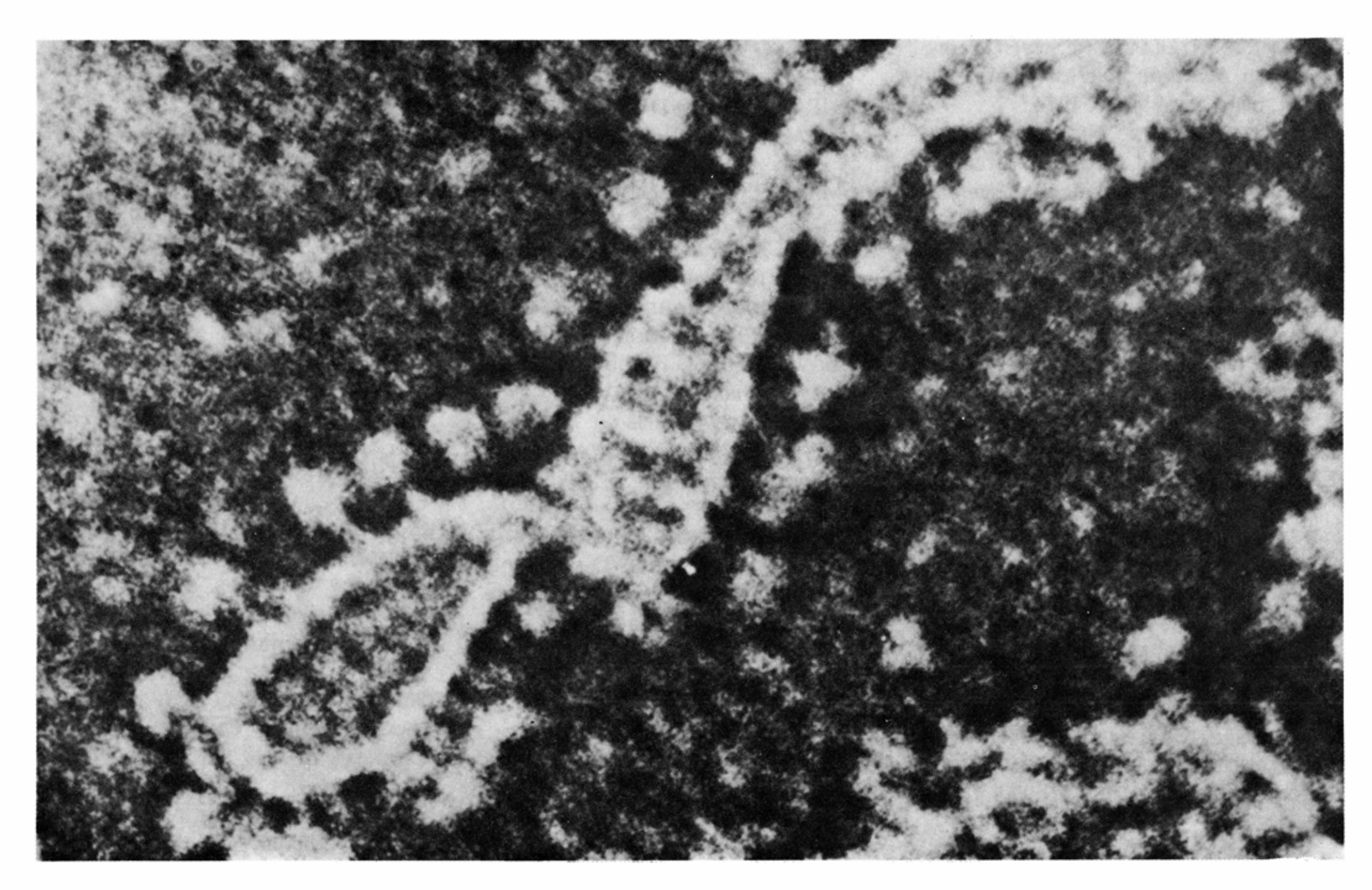
Sigue el tema de la luz y la vida. Una visión moderna, admirable, del fenómeno de la bioluminiscencia. Está llena de novedad. Ciertas proteínas incoloras, extraídas de celenterados, emiten luz centelleante cuando se les añade calcio iónico, en una proporción apreciable de 10 o 20 por ciento. La molécula almacenada no necesita ser enzimáticamente modificada para producir luz; la proteína en estado límite de excitación resplandece cuando interacciona con el ion. *In vivo,* la molécula se encierra en algún compartimiento y es accionada, de una manera muy parecida, por potenciales de acción eléctricos. Las chispas de luz que brillan en las olas y en los rizos de los mares tropicales se deben a los dinoflagelados. Las partículas celulares brillan cuando baja su *p*H en presencia de oxígeno.

Aquí la proteína portadora "precargada" libera su excitación siempre que la molécula se halla presta a alcanzar la enzima. El organismo unicelular se agita violentamente con el mar, pero la diferencia inercial entre la célula y sus flagelos flexiona aquel tenue enlace y provoca cierto potencial de acción que desencadena el resplandor.

¿Cómo puede brillar un microorganismo? Ni tiene ojos ni complementos señalizadores.

La púrpura de la visión, o rodopsina, es una proteína especializada, muy conocida, de la retina. A través de ella, la energía de los fotones excita a las neuronas. Una comunicación breve, aunque exhaustiva, del biólogo de Tubinga, profesor Dieter Oesterhelt, analiza el importante trabajo que ha llevado a cabo en unión con sus colegas de California, en los últimos años. Observaron que la púrpura de las membranas de una célula bacteriana vivía con concentraciones salinas muy elevadas. En tales halobacterias, la púrpura retiniana metaboliza produciendo ATP con energía fotónica, en lugar de hacerlo con clorofila, de la que carecen esas bacterias. El fotosensor de información se convierte allí en fototransductor energético. Hay pruebas de que el mecanismo fundamental consista en una bomba de protón portadora de luz en que se origina ATP al hacerse la luz. La bomba –situada en el interior de la estructura de la membrana purpúrea– produce gradientes en el potencial de acción y en el *p*H, que de alguna forma convierten adenosindifosfato (ADP) en ATP sin metabolitos intermedios.

Por último, la energía induce vida en movimiento. El ATP de alguna forma suministra energía a las fibras paralelas, largas e interfoliadas, del músculo estriado, deslizándose de una a otra sin cambio alguno en la longitud de la fibra. Este modelo viene analizado en profundidad por Huxley y Torkel Weis-Fogh en particular. Su estudio parte de la existencia de tres motores biológicos distintos. Uno, en la escala inferior, es el movimiento de rotación de las raíces de los flagelos de las bacterias. No es inmediatamente dependiente de ATP. El segundo es el mundo del movimiento muscular que los mamíferos comparten con otras formas avanzadas, el "desplazamiento activo" dependiente de la hidrólisis de ATP. El tercer movimiento es nuevo: un sistema genuinamente contráctil de bastoncillos, de estructuras diríanse gomosas, cuya longitud depende directamente de la concentración del ion calcio. Estos se han advertido claramente en ciertos ciliados marinos muy parecidos a pequeñas anémonas de mar.



Microfotografía electrónica, de H. Fernández-Morán, de las moléculas globulares de ATPasa alineadas a lo largo de dos mitocondrias.

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

¿EXPANSION INDEFINIDA DEL UNIVERSO?

MODERN COSMOLOGY. D. W. Sciama. Cambridge University Press, 1971.

THE AGE OF THE ELEMENTS. David N. Schramm en *Scientific American,* vol. 230, n.º 1, págs. 69-77; enero, 1974.

WHAT CAN DEUTERIUM TELL US? David N. Schramm y Robert V. Wagoner en *Physics Today,* vol. 27, n.º 12, págs. 41-47; diciembre, 1974.

AN UNBOUND UNIVERSE? J. Richard Gott III, James E. Gunn, David N. Schramm y Beatrice M. Tinsley en *The Astrophysical Journal,* vol. 194, págs. 543-553; diciembre, 1974.

LA BUSQUEDA DE NUEVAS FAMILIAS DE PARTICULAS ELEMENTALES

INTRODUCTION TO HIGH ENERGY PHYSICS. Donald H. Perkins. Addison-Wesley Publishing Co., Inc., 1972.

UNIFIED THEORIES OF ELEMENTARY-PARTICLE INTERACTION. Steven Weinberg en *Scientific American,* vol. 231, n.º 1, págs. 50-59; julio, 1974.

THE DETECTION OF NEUTRAL WEAK CURRENTS. David B. Cline, Alfred K. Mann y Carlo Rubbia en *Scientific American,* vol. 231, n.º 6, págs. 108-119; diciembre, 1974.

PROBING THE WEAK FORCE WITH NEUTRINOS. David B. Cline, Alfred K. Mann y Carlo Rubbia en *Physics Today,* vol. 28, n.º 3, págs. 23-31; marzo, 1975.

LAS PROTEINAS CROMOSOMICAS Y LA REGULACION DE LOS GENES

THE BIOLOGY OF ISOLATED CHROMATIN. James Bonner, Michael E. Dahmus, Douglas Fambrough, Ru-chih C. Huang, Keiji Marushige y Dorothy Y. H. Tuan en *Science,* vol. 159, n.º 3810, págs. 47-56; enero 5, 1968.

THE STRUCTURE AND BIOLOGICAL FUNCTION OF HISTONES. L. S. Hnilica. CRC Press, 1972.

ACIDIC PROTEINS OF THE NUCLEUS. Dirigido por Iván L. Cameron y James R. Jeter, Jr. Academic Press, 1974.

MONHISTONE CHROMOSOMAL PROTEINS AND GENE REGULATION. G. S. Stein, T. C. Spelsberg y L. J. Kleinsmith en *Science,* vol. 183, págs. 817-824; marzo 1, 1974.

CHROMOSOMAL PROTEINS AND THEIR ROLE IN THE REGULATION OF GENE EXPRESSION. Dirigido por G. S. Stein, J. S. Stein y L. J. Kleinsmith. Academic Press, en prensa.

BIOLOGIA DE LOS EMBALSES

LAKE KARIBA: A MAN-MADE TROPICAL ECOSYSTEM IN CENTRAL AFRICA. Dirigido por E. K. Balon y A. G. Coche. W. Junk Publ., The Hague, 1974.

LIMNOLOGÍA DE LOS EMBALSES ESPAÑOLES. Dept. Ecología, Univ. Barcelona, 1976. Ministerio de Obras Públicas, Madrid.

COUPLING OF LAND AND WATER SYSTEMS. Dirigido por A. D. Hasler, Springer Verlag, Berlin, 1975.

MAN-MADE LAKES AS MODIFIED ECOSYSTEMS. SCOPE Working group on man-made lakes. SCOPE Rept. 2, Scientific Committee on Problems of the Environment. International Council of Scientific Unions, Paris, 1972.

EL SURGIMIENTO DE UNA CLASE MERCANTIL MAYA

THE MAYA CHONTAL INDIANS OF ACALAN-TIXCHEL. France V. Scholes y Ralph L. Roys. University of Oklahoma Press, 1968.

THOUGHTS ON UPWARD COLLAPSE: AN ESSAY ON EXPLANATION IN ANTHROPOLOGY. Charles J. Erasmus en *Southwestern Journal of Anthropology,* vol. 24, págs. 170-194; 1968.

MAYA HISTORY AND RELIGION: J. E. S. Thompson. University of Oklahoma Press, 1970.

A STUDY OF CHANGING PRE-COLUMBIAN COMMERCIAL SYSTEMS. Dirigido por Jeremy A. Sabloff y William L. Rathje. Peabody Museum Monographs, n.º 3, 1975.

SISTEMAS ROBOT

EXPLORATORY RESEARCH IN INDUSTRIAL MODULAR ASSEMBLY: REPORT R-800. James L. Nevins, Daniel E. Whitney et al. The Charles Stark Draper Laboratory, Inc., 1974.

PROCEEDINGS OF THE FIFTH INTERNATIONAL SYMPOSIUM ON INDUSTRIAL ROBOTS: SEPTIEMBRE, 22-24, 1975. Sociedad de Ingenieros Industriales, 1975.

A NEW APPROACH TO MANIPULATOR CONTROL: THE CEREBELLAR MODEL ARTICULATION CONTROLLER (CMAC). J. S. Albus en *Transactions of the American Society of Mechanical Engineers: Series G,* Vol. 97, n.º 3, págs. 220-227; septiembre, 1975.

EL PROBLEMA DEL CANCER

END RESULTS IN CANCER: REPORT n.º 4. Dirigido por L. M. Axtell, S. J. Cutler y M. H. Myers. Publicación N. (NIH) 73-272, Departamento de Sanidad, Educación y Bienestar, 1972.

FREQUENCY OF SKIN CANCER AND SOLAR KERATOSES IN A RURAL SOUTHERN COUNTY AS DETERMINED BY POPULATION SAMPLING. Z. W. Zagula-Mally, E. W. Rosenberg y M. Kashgarian en *Cancer,* vol. 34, págs. 345-349; 1974.

IMPACT OF CANCER: YEARS OF LIFE LOST DUE TO CANCER" J. L. Murray y L. M. Axtell en *The Journal of the National Cancer Institute,* vol. 52, págs. 3-7; 1974.

VINOS, VIDES Y CLIMA

ÉLÉMENTS DE VITICULTURE GÉNÉRALE. J. Branas, G. Bernon y L. Levadoux. Montpellier, 1946.

LA SÉLÉCTION ET L'HYBRIDATION CHEZ LA VIGNE. L. Levadoux. Montpellier, 1951.

CÉPAGES ET VIGNOBLES DE FRANCE. P. Galet. Montpellier, 1956-1962.

AMERICAN WINES AND WINE-MAKING. Philip M. Wagner. Alfred A. Knopf, 1974.

JUEGOS MATEMATICOS

THE FUNDAMENTAL IDEAS OF CHINESE SCIENCE: (G) THE SYSTEM OF THE BOOK OF CHANGES en *Science and Civilisation in China:* vol. 2. Joseph Needham. Cambridge University Press, 1956.

COMBINATORIAL ASPECTS OF THE HEXAGRAMS IN THE CHINESE BOOK OF CHANGES. F. van der Blij en *Scripta Mathematica,* vol. 28, n.º 1, págs. 37-49; mayo, 1967.

I CHING, Edición, prólogo, traducción y notas de Mirko Lauer, Barral Editores, Barcelona, 1975[7].